ACCEPTED AUTHOR VERSION, TO BE PUBLISHED IN THE *JOURNAL OF EXPERIMENTAL PSYCHOLOGY: LEARNING, MEMORY, AND COGNITION.*
http://dx.doi.org/10.1037/xlm0001063

Familiarity is familiarity is familiarity:

Event-related brain potentials reveal qualitatively similar representations of personally familiar and famous faces.

Holger Wiese[1]*, Georgina Hobden[1], Eike Siilbek[1], Victoire Martignac[1], Tessa R. Flack[2], Kay L. Ritchie[2], Andrew W. Young[3], & A. Mike Burton[3]

[1]*Department of Psychology, Durham University, UK*

[2]*School of Psychology, University of Lincoln, UK*

[3]*Department of Psychology, University of York, UK*

Author Note

Parts of these findings were presented as a talk at the Virtual Vision Science Society (V-VSS) meeting 2020. We have no conflict of interest to disclose.

Correspondence concerning this aricle should be addressed to Holger Wiese, Department of Psychology, Durham University, South Road, Durham, DH1 3LE, United Kingdom. E-mail: holger.wiese@durham.ac.uk

Abstract

Humans excel in familiar face recognition, but often find it hard to make identity judgements of unfamiliar faces. Understanding of the factors underlying the substantial benefits of familiarity is at present limited, but the effect is sometimes qualified by the way in which a face is known – for example, personal acquaintance sometimes gives rise to stronger familiarity effects than exposure through the media. Given the different quality of personal versus media knowledge, for example in one's emotional response or level of interaction, some have suggested qualitative differences between representations of people known personally or from media exposure. Alternatively, observed differences could reflect quantitative differences in the level of familiarity. We present four experiments investigating potential contributory influences to face familiarity effects in which observers view pictures showing their friends, favourite celebrities, celebrities they dislike, celebrities about whom they have expressed no opinion, and their own face. Using event-related potential indices with high temporal resolution and multiple highly varied everyday ambient images as a strong test of face recognition, we focus on the N250 and the later Sustained Familiarity Effect (SFE). All known faces show qualitatively similar responses relative to unfamiliar faces. Regardless of personal- or media-based familiarity, N250 reflects robust visual representations, successively refined over increasing exposure, whilst SFE appears to reflect the amount of identity-specific semantic information known about a person. These modulations of visual and semantic representations are consistent with face recognition models which emphasise the degree of familiarity but do not distinguish between different types of familiarity.

Keywords: Face Recognition, Familiarity, Event-Related Potential, N250, Sustained Familiarity Effect

Word count: 11,656 words

Familiarity is familiarity is familiarity:

Event-related brain potentials reveal qualitatively similar representations of personally familiar and famous faces.

Human viewers are remarkably good at recognising familiar faces. We can identify people we know over a huge range of viewing conditions, even in severely degraded or distorted images (Burton, Wilson, Cowan, & Bruce, 1999; Hole, George, Eaves, & Rasek, 2002), and apparently even if we know their faces not from real life interactions but only via media exposure (Burton, Kramer, Ritchie, & Jenkins, 2016). In contrast, most of us are much poorer at recognising strangers, for example when asked to match images of someone we do not know (e.g., Bruce et al., 1999).

Accounting for the substantial and pervasive benefits of familiarity is of critical importance to understanding face recognition (Bruce & Young, 1986; Young & Burton, 2017). The benefits for familiar over unfamiliar face recognition are clearly in part perceptual in nature (Kramer, Young, & Burton, 2018), but there is also a conceptual aspect. Recognising a familiar face not only involves classifying a complex and highly variable visual stimulus (reflecting perceptual processing) but also typically allows an observer to bring to mind identity-specific semantic information (reflecting conceptual processing) that is essential to interpreting someone's behaviour and to inform socially appropriate interaction (Bruce & Young, 1986). Accessing such information to understand and guide behaviour arguably represents the most important purpose of familiar face recognition.

Identity-specific information can itself be remarkably varied and wide-ranging; including a person's occupation, past conversations, where they went on holiday, and whether we like them or not. In addition, we have different quantities of information for different people, depending on how well we know a specific person. These observations elicit questions of what exactly familiarity contributes and how access to pertinent identity-specific information is achieved. Relatively little is at present understood about these central questions (for reviews discussing related issues, see Gobbini & Haxby, 2007; Natu & O'Toole, 2011; Schweinberger & Neumann, 2016; Young & Burton, 2017). The present research examined three potential key elements, namely (i) degree and (ii) type of familiarity, as well as (iii) the integration of identity-specific, and in particular emotional information.

Psychological experiments on familiar face recognition have tended to rely on a binary contrast between 'familiar' faces (typically celebrities) and completely 'unfamiliar' faces that the participant has not seen before (such as media personalities famous only in a different country). However, it is becoming clear that familiarity is not simply a binary contrast; we are more or less familiar with particular people and we learn new faces throughout our lives. This *degree* of familiarity has perceptual consequences that can be measured and modelled (Clutterbuck & Johnston, 2002; Kramer et al., 2018). Yet it remains unclear how varying degrees of familiarity influence our mental representations.

One aspect of mental representation that has engaged the interest of researchers involves the source or *type* of familiarity, and specifically whether people are known personally or via the media (Carbon, 2008; Ramon & Gobbini, 2018). Intuitively, these types of familiarity seem quite different on a number of dimensions. Our visual exposure to personally familiar people occurs in the real 3d world, giving a broader range than our exposure to people largely known from mainly 2d media. Moreover, personally familiar people offer the option of interaction with us, something not available for media personalities. Finally, it seems that, in general, we have a larger range of affective responses to the people we know than to media stars. Our personal relationships seem to have a different quality than our attitudes towards public figures. For these reasons, some researchers have suggested that our representations of familiarity are *qualitatively* different for the people we know in person compared to those we know only via the media (Carbon, 2008; Gobbini, Leibenluft, Santiago, & Haxby, 2004; Herzmann, Schweinberger, Sommer, & Jentzsch, 2004; Sugiura et al., 2006). Evidence in favour of this suggestion has mostly come from cognitive neuroscience studies, but only a few experiments have directly compared personally familiar and famous faces. This lack of direct contrast between *type*s of familiarity in the literature has made it difficult to understand the potentially different processes involved (Natu & O'Toole, 2011).

In addition to the degree and type of familiarity, a third important aspect concerns our *emotional* response to a familiar person. As noted above, it appears that affective responses are stronger for personally familiar than famous faces (e.g., Herzmann et al., 2004). In addition to this dimension of intensity (or arousal), an emotional response can differ in valence, i.e., it can be either positive or negative. The importance of these affective components for familiarity responses has been impressively demonstrated in cases of prosopagnosia, in which affective responses in the absence of explicit recognition have been shown (Bauer, 1984; Tranel, Damasio, & Damasio, 1995), and Capgras syndrome, in which the opposite pattern has been observed (Ellis & Lewis, 2001; Ellis, Young, Quayle, & De Pauw, 1997).

A major constraint on investigating recognition of familiar faces is that performance is often so close to ceiling that it is difficult to find ways to probe the nature of familiar face representations in purely behavioural tasks (Burton et al., 2016; Clutterbuck & Johnston, 2002). A particularly promising way to address the underlying representations of familiar faces is therefore to use measures derived from human EEG (e.g., Ambrus, Kaiser, Cichy, & Kovacs, 2019; Campbell, Louw, Michniak, & Tanaka, 2020; Zimmermann, Yan, & Rossion, 2019), and especially the excellent temporal sensitivity of event-related brain potentials (ERPs). ERPs reflect voltage changes in the EEG that are time-locked to specific events, such as the presentation of visual stimuli, and consist of a number of so-called components involving distinct positive and negative deflections (e.g., Luck, 2014). While the earliest face-sensitive ERP component, the N170 (Bentin, Allison, Puce, Perez, & McCarthy, 1996), can distinguish faces from other visual objects (Eimer, 2011; Itier & Taylor, 2004; Rossion & Jacques, 2008), familiarity effects have been most consistently observed in the N250 (Andrews, Burton, Schweinberger, & Wiese, 2017; Gosling & Eimer, 2011; Kaufmann, Schweinberger, & Burton, 2009; Saavedra, Iglesias, & Olivares, 2010; Tanaka, Curran, Porterfield, & Collins, 2006), an occipito-temporal component that peaks at around 250ms and begins approximately 200ms after stimulus onset. This N250 familiarity effect is typically interpreted to reflect accessing a visual representation of a familiar face.

A more recently reported ERP response to familiar faces is the Sustained Familiarity Effect, or SFE (Wiese, Ingram, et al., 2019; Wiese, Tuttenberg, et al., 2019), observed later in time (400-600ms) at occipito-temporal electrodes. As the SFE arises later than the N250, it seems a good candidate for the involvement of processes subsequent to the initial perceptual face recognition stage, such as the integration of visual with other identity-specific information, for example factual knowledge or an emotional response to the person. Wiese and colleagues (2019) directly compared the SFE to personally familiar and celebrity faces. A clear effect was observed for the highly familiar photos of friends or relatives (by comparison to unfamiliar faces), but the SFE was absent for celebrities. However, it remains unclear whether this pattern reflects a qualitative difference due to *types* of familiarity (celebrity versus personally familiar) or a quantitative difference due to *degree* of familiarity (as personally familiar faces were likely more familiar).

To investigate the effects of *degree* and *type* of familiarity, Experiments 1 and 2 therefore measured the N250 and SFE to images of faces of personal friends or relatives and favourite celebrities (known only through the media, but, like friends, very familiar and of positive valence) in comparison to different types of unfamiliar faces. Experiment 3 then manipulated the *emotional* dimension of valence by using faces of liked (positive valence) or disliked (negative valence) celebrities. Finally, Experiment 4 probed the nature and extent of personal knowledge by recording the N250 and SFE for the participants' own, relative to a personally familiar face. Across all four experiments multiple and highly varied ambient images of faces sampled from the real world were used as stimuli (cf. Jenkins & Burton, 2011), rather than controlled posed photographs taken specifically for research purposes, so that effects involving the N250 and SFE would be demonstrably robust across image differences and hence could be attributed to mechanisms involved in everyday recognition.

Experiment 1: Personally familiar faces versus favourite celebrities

Experiment 1 was designed to test whether personally familiar faces are processed qualitatively differently than celebrity faces. As noted above, previous work did not detect an SFE (or N250) to celebrity faces (Wiese, Tuttenberg, et al., 2019). This absence of an SFE might suggest that famous and personally familiar faces are indeed processed differently, which in turn might point to a different *type* of representation for celebrities. Alternatively, however, celebrities in this previous study might have not been sufficiently familiar, and the absence of an SFE might therefore reflect an insufficient *degree* of familiarity suggesting a *quantitative* difference.

To tease apart these possibilities, Experiment 1 presented multiple ambient images of the participants' favourite celebrities, ensuring a high degree of familiarity. The absence of an SFE for favourite celebrities would then point more clearly to a qualitative difference between the representation of media-based versus personal familiarity. Personally familiar faces of participants' friends or relatives were also presented in Experiment 1, to enable a direct comparison of potentially different familiarity types. Finally, unfamiliar celebrities (e.g., singers or actors only known in countries other than the UK) as well as non-famous unfamiliar faces were presented. This use of two different types of unfamiliar face was included to estimate any influences of potential systematic differences that might exist between pictures of celebrities and non-famous people that are not directly related to familiarity (e.g., differences between image sets due to professional photography or make-up in celebrity pictures, or differences in average attractiveness or distinctiveness).

We expected clear ERP familiarity effects, both in the N250 and SFE time range, for personally familiar versus unfamiliar faces. Crucially, if celebrity faces are indeed represented in a qualitatively different way, we predicted no SFE for favourite celebrities. If, however, quantitative differences in the degree familiarity explained our previous finding, we expected a clear SFE in this condition. Finally, as some previous studies observed familiarity effects in the N170 (e.g., Caharel, Courtay, Bernard, Lalonde, & Rebai, 2005; Caharel, Jacques, d'Arripe, Ramon, & Rossion, 2011; Johnston, Overell, Kaufman, Robinson, & Young, 2016), we additionally analysed this component.

Methods

Participants

The necessary sample size was estimated by conducting a power analysis using G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007). Assuming half the SFE effect size for favourite celebrities than for personally familiar faces in our previous experiment (Experiment 3 in Wiese, Tuttenberg, et al., 2019), this analysis suggested a total N of 19 to detect significantly more negative amplitudes for favourite celebrities than unfamiliar faces (paired-sample t-test, one-tailed, $dz = 0.8$, $1 - \beta = .95$). Twenty-two Durham University undergraduate students were tested, two of whom were excluded due to technical problems during EEG recordings. The final sample of 20 participants consisted of 18 females and two males with a mean age of 19.9 years (SD = 1.1). All participants received course credit or monetary compensation, were right-handed according to a modified version of the Edinburgh Handedness Inventory (Oldfield, 1971), reported normal or corrected-to-normal vision, and did not take central-acting medication. All participants gave written informed consent, and the experiment was approved by the ethics committee of Durham University's Department of Psychology.

Stimuli

Each participant provided 50 digital images of a highly personally familiar face (close friends or relatives) and 50 images of their favourite celebrity (e.g. favourite actor, singer, athlete etc.). Consent of the depicted non-famous people was obtained via email. Moreover, 50 images of an unfamiliar non-famous identity (ID) and 50 images of an unfamiliar famous ID (i.e., actors or singers from other countries, who were unknown to the participants) were chosen from a set of ten non-famous unfamiliar and nine unfamiliar celebrity IDs to match basic characteristics of the familiar IDs (gender, ethnicity, approximate hair colour). Eight different images of butterflies were used as targets in the ERP paradigm. Rectangles around the faces and butterflies were cropped from the original images, resized, copied into a frame of 190 x 285 pixels, and converted to grey scale (see Figure 1a). Images were matched for luminance using the SHINE toolbox (Willenbockel et al., 2010).

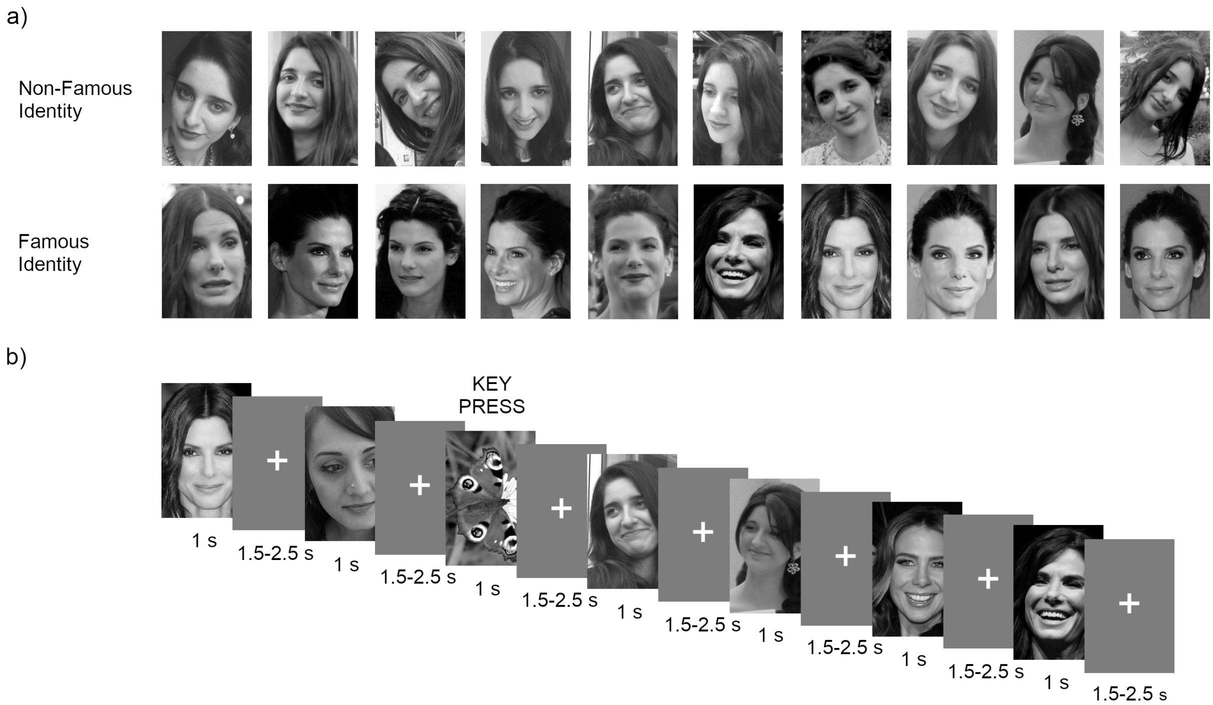

Figure 1. a) Ambient images of non-famous and famous identities, similar to those used in Experiments 1-4. Images of the celebrities are reproduced here under creative commons licensing (see supplementary material for full copyright information for each image). Images of the non-famous identities are reproduced here with their consent. b) Trial structure of the experiments.

Procedure

Participants were seated in an electrically shielded chamber (Global EMC^{TM}) with their heads in a chin rest 80 cm from a computer monitor. The experiment consisted of a single block of 220 trials, in which all 50 images of the four respective facial IDs were presented once, in random order, intermixed with 20 trials showing butterflies. Each image was presented at a visual angle of 3.6° × 5.4° for 1,000 ms, followed by a fixation cross for between 1,500 ms and 2,500 ms (2,000 ms on average). Participants were instructed to pay attention to the screen and to respond with a right index finger button press as quickly as possible whenever a butterfly was presented (see Figure 1b).

Following the main experiment, participants were presented with eight randomly selected images shown simultaneously on the screen of each of the four IDs used in their respective version of the experiment. Participants were asked to rate each ID for familiarity (“How likely is it that you would recognise this person?”, from 1 = very unlikely to 5 = very likely), valence ("How do you feel when you see this person?", from 1 = very positive to 5 = very negative), and arousal ("How do feel when you see this person?", from 1 = very excited to 5 = not excited at all). Valence and arousal ratings were illustrated using the Self-Assessment Mannequin scale (Bradley & Lang, 1994).

EEG recording and data analysis

During the main experiment, 64-channel EEG was recorded from DC to 200 Hz, with a 1024 Hz sampling frequency using sintered Ag/Ag-Cl electrodes (EEGo, ANT Neuro, Hengelo, The Netherlands). An electrode on the forehead (AFz) served as ground, and CPz was used as the recording reference. Blinks were corrected using the algorithm implemented in BESA Research software (Version 6.3; Gräfelfing, Germany). The EEG was then segmented from -200 to 1,000 ms relative to stimulus onset, with the first 200 ms as the baseline. Artifact rejection was carried out using an amplitude threshold of 100 µV and a gradient criterion of 75 µV, and trials with incorrect button presses were excluded. Remaining trials were recalculated to the common average reference and averaged separately for each experimental condition. Average numbers of analysed trials were 48.2 (SD = 3.2, min = 39) for personally familiar faces, 48.2 (SD = 3.8, min = 37) for favourite celebrities, 47.5 (SD = 3.8, min = 38) for unfamiliar celebrities, and 48.2 (SD = 3.7, min = 38) for unfamiliar faces.

Mean amplitudes from 140 to 180 ms (N170), from 200 to 400 ms (N250), and from 400 to 600 ms (SFE) were analysed at occipito-temporal electrodes TP9/TP10 and P9/P10 using Analyses of Variance (ANOVA) and t-tests. Additional Bayesian tests on selected pairwise comparisons are reported in a supplement. These time windows for the N250 and SFE are identical to those in our previous paper (Wiese, Tuttenberg, et al., 2019) and were chosen prior to data analysis. Similarly, electrodes TP9/TP10 were used in our previous paper as the electrodes of interest and were also chosen before data analysis. In addition, and again prior to data analysis, we decided to include the neighbouring electrodes P9/P10 to not miss potentially more posterior distributions of any familiarity effect we examined here (see also Wiese, Ingram, et al., 2019). These decisions corresponded well with the distribution and timing of ERPs in this experiment. Additional analyses involving all electrodes are reported in a supplement.

Following an estimation approach (Cumming, 2012), we report effect size measures with appropriate confidence intervals (CIs) throughout (e.g., Fritz, Morris, & Richler, 2012). CIs for partial eta squared were calculated using scripts provided by M.J. Smithson (www.michaelsmithson.online/stats/CIstuff/CI.html). Cohen's d for repeated-measures was bias-corrected and calculated using ESCI (Cumming & Calin-Jageman, 2017), with the mean standard deviation rather than the standard deviation of the difference as the denominator. In addition, the reliability of ERP familiarity effects in individual participants was calculated using a bootstrapping technique (Di Nocera & Ferlazzo, 2000) involving 10,000 random re-assignments of individual participants' single-trial EEG epochs to the unfamiliar versus one of the other three conditions (i.e., personally familiar, favourite celebrity or unfamiliar celebrity) in separate analyses. We assumed a reliable effect if the true individual effect was larger than 95% of random re-samplings. To keep these analyses comparable to our previous studies, we conducted them at electrodes TP9/TP10. Data is available in a publicly accessible repository (https://osf.io/4me6v/?view_only=4aa5013f128a422eac3f804df18fe99f).

Results

Performance

Participants performed at near ceiling level during the butterfly detection task, hit rate = .98, SD = .04, false alarm rate < .01, SD = .01, showing that they were attentive to the stimuli throughout the experiment. Mean reaction time for correct responses was 505 ms, SD = 79.

EEG/Event-related potentials

Visual inspection of the ERP waveforms suggested clear familiarity effects for both personally familiar and favourite celebrity faces from approximately 200 ms after stimulus onset (see Figure 2). A repeated-measures ANOVA in the N170 time range (140-180 ms) with the within-subjects factors hemisphere (left, right), site (TP, P), identity type (famous, non-famous) and familiarity (familiar, unfamiliar) revealed an interaction of site by identity type, $F(1, 19) = 4.65$, $p = .044$, $\eta^2_p = .197$, 90% CI [.003, .416], with more negative amplitudes for famous relative to non-famous faces particularly at more anterior electrode positions. No significant effects involving the familiarity factor were detected at this early latency, all $F < 2.69$, all $p > .118$, all $\eta^2_p < .124$.

A corresponding ANOVA in the N250 time range (200-400 ms) revealed a significant main effect of familiarity, $F(1, 19) = 46.59$, $p < .001$, $\eta^2_p = .710$, 90% CI [.471, .802], as well as a significant interaction of familiarity by identity type, $F(1, 19) = 7.53$, $p = .013$, $\eta^2_p = .284$, 90% CI [.038, .492]. The N250 was significantly more negative for personally familiar relative to unfamiliar faces, $M_{diff.} = 2.46$ µV, 95% CI [1.67, 3.25], $t(19) = 6.53$, $p < .001$, $d_{unb.} = 0.52$, 95% CI [0.30, 0.77], and for favourite celebrities versus unfamiliar faces, $M_{diff.} = 1.88$ µV, 95% CI [0.93, 2.84], $t(19) = 4.13$, $p = .001$, $d_{unb.} = 0.42$, 95% CI [0.18, 0.68]. There were no significant differences between unfamiliar celebrities (i.e., celebrities famous in different countries) and unfamiliar faces, $M_{diff.} = 0.70$ µV, 95% CI [-0.38, 1.77], $t(19) = 1.36$, $p = .189$, $d_{unb.} = 0.14$, 95% CI [-0.07, 0.36], or personally familiar versus favourite celebrity faces, $M_{diff.} = 0.57$ µV, 95% CI [-0.22, 1.37], $t(19) = 1.51$, $p = .147$, $d_{unb.} = 0.12$, 95% CI [-0.05, 0.30]. No other effects involving the identity type factor were significant, all $F < 3.80$, all $p > .065$, all $\eta^2_p < .167$.

Following a reviewer suggestion, we additionally calculated analyses using vector-length based corrections (McCarthy & Wood, 1985; see also Urbach & Kutas, 2002) to examine potential differences in scalp distribution for different types of familiarity. An interaction of hemisphere by familiarity by identity type was observed, $F(1, 19) = 6.60$, $p = .019$, $\eta^2_p = .258$, 90% CI [.026, .470]. Follow up tests revealed hemisphere effects for all four face categories tested, all $F > 10.78$, all $p < .005$, all $\eta^2_p > .361$, suggesting no difference in scalp distribution.

An ANOVA in the SFE time range (400-600 ms) again revealed a significant main effect of familiarity, $F(1, 19) = 111.45$, $p < .001$, $\eta^2_p = .854$, 90% CI [.714, .900], as well as an interaction of familiarity by hemisphere, $F(1, 19) = 6.00$, $p = .024$, $\eta^2_p = .240$, 90% CI [.018, .455]. The familiarity effect (i.e. the difference between familiar and unfamiliar faces) was significantly larger over the right hemisphere for personally familiar faces (hemisphere by familiarity interaction: $F[1, 19] = 7.31$, $p = .014$, $\eta_p^2 = .278$, 90% CI [.035, 487]), while the corresponding effect for favourite celebrities was observed as a statistical trend ($F[1, 19] = 4.08$, $p = .058$, $\eta_p^2 = .177$, 90% CI [.000, .398]). Moreover, in the omnibus ANOVA, the interaction of familiarity by identity type was significant, $F(1, 19) = 10.86$, $p = .004$, $\eta^2_p = .364$, 90% CI [.085, .555]. No other effects involving the identity type factor were significant, all $F < 3.74$, all $p > .067$, all $\eta^2_p < .163$. Planned contrasts yielded more negative amplitudes for personally familiar versus unfamiliar faces, $M_{diff.} = 3.71$ µV, 95% CI [2.80, 4.61], $t(19) = 8.55$, $p < .001$, $d_{unb.} = 0.97$, 95% CI [0.61, 1.40], and for favourite celebrities versus unfamiliar faces, $M_{diff.} = 2.28$ µV, 95% CI [1.50, 3.05], $t(19) = 6.15$, $p < .001$, $d_{unb.} = 0.66$, 95% CI [0.38, 0.99], but not for unfamiliar celebrities versus unfamiliar faces, $M_{diff.} = 0.59$ µV, 95% CI [-0.47, 1.65], $t(19) = 1.17$, $p = .257$, $d_{unb.} = 0.15$, 95% CI [-0.11, 0.43]. Personally familiar faces elicited more negative amplitudes relative to favourite celebrities, $M_{diff.} = 1.43$ µV, 95% CI [0.39, 2.46], $t(19) = 2.88$, $p = .010$, $d_{unb.} = 0.38$, 95% CI [0.09, 0.68]. A further ANOVA using vector-length based correction to test for potential differences in scalp distribution did not reveal any significant interaction of experimental conditions with hemisphere or site factors, all $F < 1.91$, all $p > .183$, all $\eta^2_p < .092$.

Finally, again following a reviewer suggestion, to test for effects independent of the earlier N250 time range, we calculated additional analyses for the SFE time window using the N250 rather than the pre-stimulus interval as the baseline. A repeated-measures ANOVA yielded significant main effects of familiarity, $F(1, 19) = 19.34$, $p < .001$, $\eta^2_p = .504$, 90% CI [.206, .659], and identity type, $F(1, 19) = 6.13$, $p = .023$, $\eta^2_p = .244$, 90% CI [.020, .459], which was further qualified by an interaction of site by identity type, $F(1, 19) = 5.06$, $p = .037$, $\eta^2_p = .210$, 90% CI [.008, .429], reflecting less negative amplitudes for famous relative to non-famous faces at TP9/10, $M_{diff.} = 0.62$ µV, 95% CI [0.26, 0.98], $t(19) = 3.57$, $p = .002$, $d_{unb}. = 0.31$, 95% CI [0.12, 0.53], but not at P9/10 sites, $M_{diff.} = 0.34$ µV, 95% CI [-0.14, 0.82], $t(19) = 1.48$, $p = .155$, $d_{unb}. = 0.13$, 95% CI [-0.05, 0.32]. Planned comparisons revealed more negative amplitudes for personally familiar faces relative to both unfamiliar celebrities, $M_{diff.} = 1.25$ µV, 95% CI [0.52, 1.97], $t(19) = 3.61$, $p = .002$, $d_{unb}. = 0.53$, 95% CI [0.20, 0.89], and unfamiliar non-famous faces, $M_{diff.} = 1.35$ µV, 95% CI [0.68, 2.03], $t(19) = 4.20$, $p < .001$, $d_{unb}. = 0.61$, 95% CI [0.27, 0.99]. Favourite celebrities were significantly more negative than unfamiliar celebrities, $M_{diff.} = 0.50$ µV, 95% CI [0.08, 0.92], $t(19) = 2.51$, $p = .021$, $d_{unb}. = 0.22$, 95% CI [0.03, 0.42], while a trend was detected for the comparison to unfamiliar non-famous faces, $M_{diff.} = 0.39$ µV, 95% CI [-0.07, 0.86], $t(19) = 1.77$, $p = .093$, $d_{unb}. = 0.16$, 95% CI [-0.03, 0.36]. Finally, personally familiar faces were more negative than favourite celebrities, $M_{diff.} = 0.85$ µV, 95% CI [0.26, 1.45], $t(19) = 2.99$, $p = .008$, $d_{unb}. = 0.37$, 95% CI [0.10, 0.66].

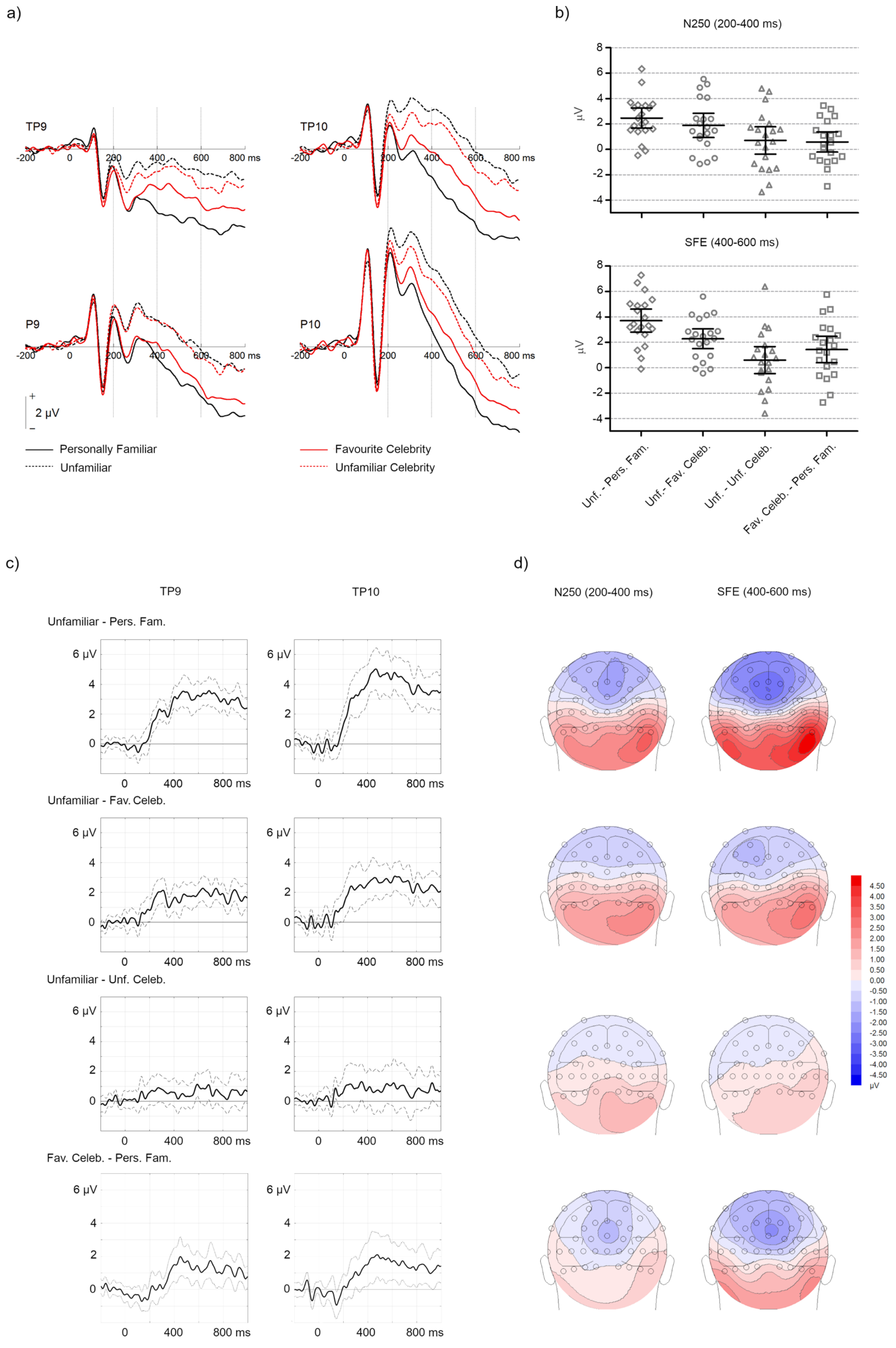

Figure 2. a) Grand average event-related potentials at left and right-hemispheric occipito-temporal electrodes TP9/TP10 and P9/P10. Dashed lines mark the N250 (200-400 ms) and SFE (400-600 ms) time ranges. b) Mean (+/- 95% CI) and individual familiarity effects in the N250 and SFE time ranges at electrodes TP9/TP10/P9/P10. c) Mean (+/- 95% CI) difference waves at left and right occipito-temporal electrodes TP9/TP10. d) Scalp-topographical voltage maps (spherical spline interpolation, 90° equidistant projection) of familiarity effects in the N250 and SFE time window.

Bootstrapping analyses in the N250 time window yielded reliable effects in 17/20 participants for personally familiar versus unfamiliar faces, Proportion (P) = .85, 95% CI [.64, .95], in 12/20 participants for favourite celebrities versus unfamiliar faces, P = .60, 95% CI [.39, .78], in 8/20 participants for unfamiliar celebrities versus unfamiliar faces, P = .40, 95% CI [.22, .61], and in 7/20 participants for personally familiar versus favourite celebrity faces, P = .35, 95% CI [.18, .57]. In the SFE time window, bootstrapping revealed reliable effects in 17/20 participants for the personally familiar versus unfamiliar face IDs, P = .85, 95% CI [.64, .95], in 12/20 participants for favourite celebrities versus unfamiliar faces, P = .60, 95% CI [.39, .78], in 5/20 participants for unfamiliar celebrities versus unfamiliar faces, P = .25, 95% CI [.11, .47], and in 10/20 participants for personally familiar versus favourite celebrity faces, P = .50, 95% CI [.30, .70].

Rating Task

The results of the rating task are reported in Table 1. Rated familiarity was significantly higher for personally familiar versus unfamiliar, $t(19) = 11.17$, $p < .001$, $d_{unb.} = 3.39$[1], and for favourite celebrities versus unfamiliar faces, $t(19) = 10.17$, $p < .001$, $d_{unb.} = 3.14$. Unfamiliar celebrities and unfamiliar faces did not differ, $t(19) = 0.59$, $p = .560$, $d_{unb.} = 0.16$, 95% CI [-0.39, 0.72], and neither did personally familiar and favourite celebrities, $t(19) = 1.45$, $p = .163$, $d_{unb.} = 0.44$, 95% CI [-0.18, 1.09].

Valence ratings were significantly more positive for personally familiar versus unfamiliar, $t(19) = 19.62$, $p < .001$, $d_{unb.} = 5.61$, and for favourite celebrities versus unfamiliar faces, $t(19) = 10.47$, $p < .001$, $d_{unb.} = 3.18$. Unfamiliar celebrities and unfamiliar faces did not differ, $t(19) = 0.30$, $p = .772$, $d_{unb.} = 0.09$, 95% CI [-0.53, 0.72]. Personally familiar faces were rated as more positive than favourite celebrities, $t(19) = 2.99$, $p = .008$, $d_{unb.} = 0.91$, 95% CI [0.25, 1.62].

[1] Note that ESCI only calculates CIs for estimates of δ between -2 and 2 (see Cumming, 2012, p. 306-307).

Finally, arousal was significantly higher for personally familiar than unfamiliar, $t(19) = 7.78$, $p < .001$, $d_{unb.} = 2.02$, and for favourite celebrities than unfamiliar faces, $t(19) = 8.39$, $p < .001$, $d_{unb.} = 2.23$. There was a trend for higher arousal for unfamiliar celebrities relative to unfamiliar faces, $t(19) = 1.99$, $p = .061$, $d_{unb.} = 0.49$, 95% CI [-0.02, 1.04]. Personally familiar and favourite celebrities did not differ, $t(19) = 0.51$, $p = .614$, $d_{unb.} = 0.13$, 95% CI [-0.39, 0.67].

		Familiarity	Valence	Arousal
Experiment 1				
	Personally Familiar	5.00 ± 0.00	1.10 ± 0.31	2.15 ± 1.23
	Favourite Celebrity	4.90 ± 0.31	1.50 ± 0.51	2.30 ± 0.92
	Unfamiliar Celebrity	2.50 ± 1.32	2.95 ± 0.69	3.85 ± 1.23
	Unfamiliar	2.30 ± 1.08	2.90 ± 0.31	4.40 ± 0.88
Experiment 2				
	Personally Familiar	5.00 ± 0.00	1.11 ± 0.32	1.84 ± 0.76
	Favourite Celebrity	4.95 ± 0.23	1.53 ±0.61	2.16 ± 0.69
	Other Celebrity	3.32 ± 1.38	2.79 ± 1.03	3.37 ± 1.26
	Unfamiliar	2.26 ± 1.45	2.95 ± 0.71	4.32 ± 0.82
Experiment 3				
	Favourite Celebrity	4.92 ± 0.28	1.36 ± 0.57	2.20 ± 0.91
	Unfamiliar Celebrity	1.36 ± 0.64	2.64 ± 0.70	3.80 ± 1.19
	Disliked Celebrity	4.72 ± 0.68	4.20 ± 0.71	2.44 ± 1.04
	Unfamiliar	1.80 ± 1.22	2.80 ± 0.71	4.04 ± 0.98
Experiment 4				
	Own Face	5.00 ± 0.00	2.00 ± 0.97	2.45 ± 1.15
	Personally Familiar	4.90 ± 0.45	1.35 ± 0.93	1.75 ± 0.97
	Other Face	2.15 ± 1.66	2.60 ± 0.88	3.50 ± 1.24
	Unfamiliar	1.60 ± 1.23	2.55 ± 0.60	3.75 ± 1.16

Table 1. Mean (+/- SD) ratings from Experiments 1 – 4. All ratings were made on a scale from 1 to 5 (familiarity: 1 = very low familiarity to 5 = very high familiarity; valence: 1 = very positive to 5 = very negative; arousal: 1 = very excited to 5 = not excited at all)

Discussion

Experiment 1 provided clear evidence for ERP familiarity effects in the N250 and SFE time ranges for both personally familiar faces and, crucially, for celebrity faces perceived as particularly well-known and liked. These findings demonstrate that media-based familiarity is sufficient to elicit the SFE, and therefore suggest qualitatively similar representations for celebrity and personally familiar faces. Interestingly, while the amplitude of the N250 did not differ, the SFE elicited a more graded pattern, with more negative waveforms for personally familiar relative to favourite celebrity faces. This graded effect was not paralleled by corresponding differences in familiarity ratings to the face images, which potentially highlight the visual aspect of recognition. The SFE, however, more likely reflects the integration of visual with additional identity-specific information. As participants will have had access to more semantic, episodic and affective information for personally familiar relative to favourite celebrity identities, the SFE might have captured this difference.

Although statistical comparisons at the group level did not reach significance, unfamiliar celebrity faces elicited reliably more negative amplitudes than unfamiliar non-famous faces in a minority of participants (40% for the N250, 25% for the SFE). As noted above, these "false familiarity effects" might reflect systematic differences between the pictures of famous and non-famous unfamiliar faces. For instance, unfamiliar celebrities may on average be more attractive and/or more distinctive than non-famous faces. Therefore, participants presumably found unfamiliar celebrities more interesting, and were more likely to learn them during the experiment. Rating data show a trend for higher arousal for unfamiliar celebrities relative to non-famous unfamiliar faces, which might be seen as supporting this suggestion. In addition, different pictures of unfamiliar celebrities may be more similar, e.g. due to similar professional photo poses, emotional expressions, and make-up, which again might have helped to learn the faces. Previous studies have shown that learning during an experiment results in more negative N250 components (Kaufmann et al., 2009; Tanaka et al., 2006), which may well explain the minor effects observed here.

Of interest in this context, ERP analyses show a main effect of "identity type" in the N170 but not in the later N250 and SFE that are of primary interest to our study. The only exception to this overall finding was a significant interaction of identity type by electrode site for the N250-corrected SFE. Crucially, however, clear familiarity effects were detected *within* each identity category, i.e. for favourite versus unfamiliar celebrities and for personally familiar versus non-famous unfamiliar faces. These effects are unlikely to be driven by image characteristics unrelated to familiarity, and it therefore appears appropriate to conclude that an SFE is detectable for both famous and non-famous faces.

Experiment 1 thus provided initial evidence for an SFE elicited by multiple ambient images of celebrity faces, and therefore qualitatively similar representations for personal and media-based familiarity. However, as our previous findings (Wiese, Tuttenberg, et al., 2019) did not suggest familiarity effects for celebrity faces, a single positive finding is not sufficient for any strong conclusions. Experiment 2 therefore aimed at replicating the basic finding of Experiment 1 in a further experiment that took a closer look at the influence of celebrity faces.

Experiment 2: Favourite versus other celebrities

Experiment 2 was designed to further investigate ERP familiarity effects for favourite celebrity and personally familiar faces. To balance image properties across experimental conditions more rigorously than the informal matching of general visual properties (gender, ethnicity, hair colour) used in Experiment 1, participants were assigned to arbitrary pairings in Experiment 2, and within each pair the personally familiar and favourite celebrity identities of participant 1 were used as the unfamiliar and "other" celebrity identities for participant 2, and vice versa. As "other" celebrities in this design were likely familiar to some extent (given that the identities were chosen by participants of the same age group and with a similar educational background), this procedure also allowed us to examine whether any potential ERP effects observed for favourite celebrities would be found for this other celebrity category.

On the basis of Experiment 1, we expected clear N250 familiarity effects and SFEs for both personally familiar and favourite celebrity relative to unfamiliar faces. Moreover, given that other celebrities were likely visually familiar to some extent, we expected to find an N250 effect for this condition relative to unfamiliar faces. However, given that participants presumably had no strong affective response and/or extensive identity-specific knowledge, we predicted no clear SFE in this condition.

Methods

Participants

Twenty-one Durham University undergraduate students were tested, one of whom was excluded due to technical problems during EEG recording. The final sample consisted of 14 females and six males, with a mean age of 20.2 years, SD = 0.8. Reimbursement and inclusion/exclusion criteria were identical to Experiment 1. All participants gave written informed consent and the experiment was approved by the ethics committee at Durham University's Psychology department.

Stimuli, procedure, EEG recording and data analysis

As for Experiment 1, each participant provided 50 images of a personally familiar ID and of their favourite celebrity, respectively. Participants were paired, and personally familiar and favourite celebrity IDs for one participant in each pair were used as the unfamiliar and other celebrity IDs for the other participant. All other aspects of the experiment, including EEG recording and data analysis parameters remained unchanged. Average numbers of trials analysed were 47.2 (SD = 2.6, min = 41) for personally familiar faces, 46.6 (SD = 3.9, min = 36) for favourite celebrities, 45.7 (SD = 4.1, min = 36) for other celebrities, and 45.8 (SD = 4.4, min = 33) for unfamiliar faces.

Results

Performance

Similar to Experiment 1, performance during the butterfly task was very accurate, mean hit rate = .97, SD = .06, mean false alarm rate = .05, SD = .22. Mean reaction time for correct responses was 561 ms, SD = 84.

Event-related potentials

Visual inspection of the ERP waveforms again suggested clear familiarity effects for both personally familiar and favourite celebrity faces from approximately 200 ms after stimulus onset (see Figure 3). A repeated-measures ANOVA in the N170 time range yielded a significant main effect of familiarity, $F(3, 57) = 4.68$, $p = .005$, $\eta^2_p = .198$, 90% CI [.039, .311]. Relative to the unfamiliar condition, favourite celebrities were significantly more negative, $M_{diff.} = 0.80$ µV, 95% CI [0.27, 1.34], $t(19) = 3.13$, $p = .006$, $d_{unb.} = 0.27$, 95% CI [0.08, 0.47], and a corresponding trend was observed in the other celebrity condition, $M_{diff.} = 0.55$ µV, 95% CI [-0.02, 1.12], $t(19) = 2.03$, $p = .057$, $d_{unb.} = 0.18$, 95% CI [-0.01, 0.37]. By contrast, personally familiar faces were not significantly different from unfamiliar faces, $M_{diff.} = 0.16$ µV, 95% CI [-0.41, 0.73], $t(19) = 0.60$, $p = .558$, $d_{unb.} = 0.06$, 95% CI [-0.14, 0.25]. At the same time, favourite celebrities elicited significantly more negative amplitudes than personally familiar faces, $M_{diff.} = 0.97$ µV, 95% CI [0.20, 1.73], $t(19) = 2.65$, $p = .016$, $d_{unb.} = 0.33$, 95% CI [0.06, 0.62].

A corresponding ANOVA in the N250 time range revealed a significant main effect of familiarity, $F(3, 57) = 10.06$, $p < .001$, $\eta^2_p = .346$, 90% CI [.158, .457]. Relative to the unfamiliar condition, planned comparisons yielded significantly more negative N250 amplitudes for personally familiar, $M_{diff.} = 1.60$ µV, 95% CI [1.03, 2.17], $t(19) = 5.82$, $p < .001$, $d_{unb.} = 0.61$, 95% CI [0.34, 0.92], favourite celebrity, $M_{diff.} = 1.61$ µV, 95% CI [0.90, 2.31], $t(19) = 4.76$, $p < .001$, $d_{unb.} = 0.57$, 95% CI [0.28, 0.89], and other celebrity faces, $M_{diff.} = 0.78$ µV, 95% CI [0.12, 1.43], $t(19) = 2.48$, $p = .023$, $d_{unb.} = 0.27$, 95% CI [0.04, 0.52]. Personally familiar and favourite celebrity faces did not differ, $M_{diff.} = 0.01$ µV, 95% CI [-0.92, 0.93], $t(19) = 0.01$, $p = .989$, $d_{unb.} < 0.01$, 95% CI [-0.32, 0.32]. An ANOVA using vector-length based correction revealed no significant interaction of familiarity with hemisphere or site factors, all $F < 1.26$, all $p > .299$, all $\eta^2_p < .063$.

Analysis of the SFE time range again revealed a significant main effect of familiarity, $F(3, 57) = 15.88$, $p < .001$, $\eta^2_p = .455$, 90% CI [.269, .554]. Planned contrasts yielded significantly more negative amplitudes for personally familiar relative to unfamiliar faces, $M_{diff.} = 2.65$ µV, 95% CI [1.69, 3.60], $t(19) = 5.80$, $p < .001$, $d_{unb.} = 0.91$, 95% CI [0.50, 1.38], as well as for favourite celebrities versus unfamiliar faces, $M_{diff.} = 2.39$ µV, 95% CI [1.42, 3.37], $t(19) = 5.16$, $p < .001$, $d_{unb.} = 0.81$, 95% CI [0.42, 1.25]. Moreover, other celebrities elicited more negative amplitudes than unfamiliar faces, $M_{diff.} = 0.89$ µV, 95% CI [0.15, 1.63], $t(19) = 2.53$, $p = .020$, $d_{unb.} = 0.33$, 95% CI [0.05, 0.63]. Personally familiar and favourite celebrity faces did not differ, $M_{diff.} = 0.25$ µV, 95% CI [-0.87, 1.38], $t(19) = 0.47$, $p = .643$, $d_{unb.} = 0.08$, 95% CI [-0.27, 0.43]. An additional ANOVA using vector-length based correction revealed a significant interaction of familiarity by hemisphere, $F(3, 57) = 3.70$, $p = .017$, $\eta^2_p = .163$, 90% CI [.018, .274]. Follow-up tests revealed a significant hemisphere effect for unfamiliar faces, $F(1, 19) = 6.07$, $p = .023$, $\eta^2_p = .242$, 90% CI [.019, .457], but not for any of the other three conditions, all $F < 1.37$, all $p > .257$, all $\eta^2_p < .068$.

A further ANOVA in the SFE time range, using the N250 rather than the pre-stimulus interval as the baseline, again yielded a significant main effect of familiarity, $F(3, 57) = 5.24$, $p = .003$, $\eta^2_p = .216$, 90% CI [.051, .331]. Planned comparisons again revealed significantly more negative amplitudes for personally familiar relative to unfamiliar faces, $M_{diff.} = 1.05$ μV, 95% CI [0.29, 1.81], $t(19) = 2.90$, $p = .009$, $d_{unb.} = 0.54$, 95% CI [0.14, 0.98], and for favourite celebrities versus unfamiliar faces, $M_{diff.} = 0.79$ μV, 95% CI [0.26, 1.32], $t(19) = 3.14$, $p = .005$, $d_{unb.} = 0.43$, 95% CI [0.13, 0.76]. However, other celebrities did not elicit more negative amplitudes than unfamiliar faces, $M_{diff.} = 0.11$ μV, 95% CI [-0.26, 0.49], $t(19) = 0.63$, $p = .536$, $d_{unb.} = 0.06$, 95% CI [-0.14, 0.27]. Finally, personally familiar and favourite celebrity faces again did not differ, $M_{diff.} = 0.26$ μV, 95% CI [-0.54, 1.06], $t(19) = 0.68$, $p = .503$, $d_{unb.} = 0.14$, 95% CI [-0.27, 0.56].

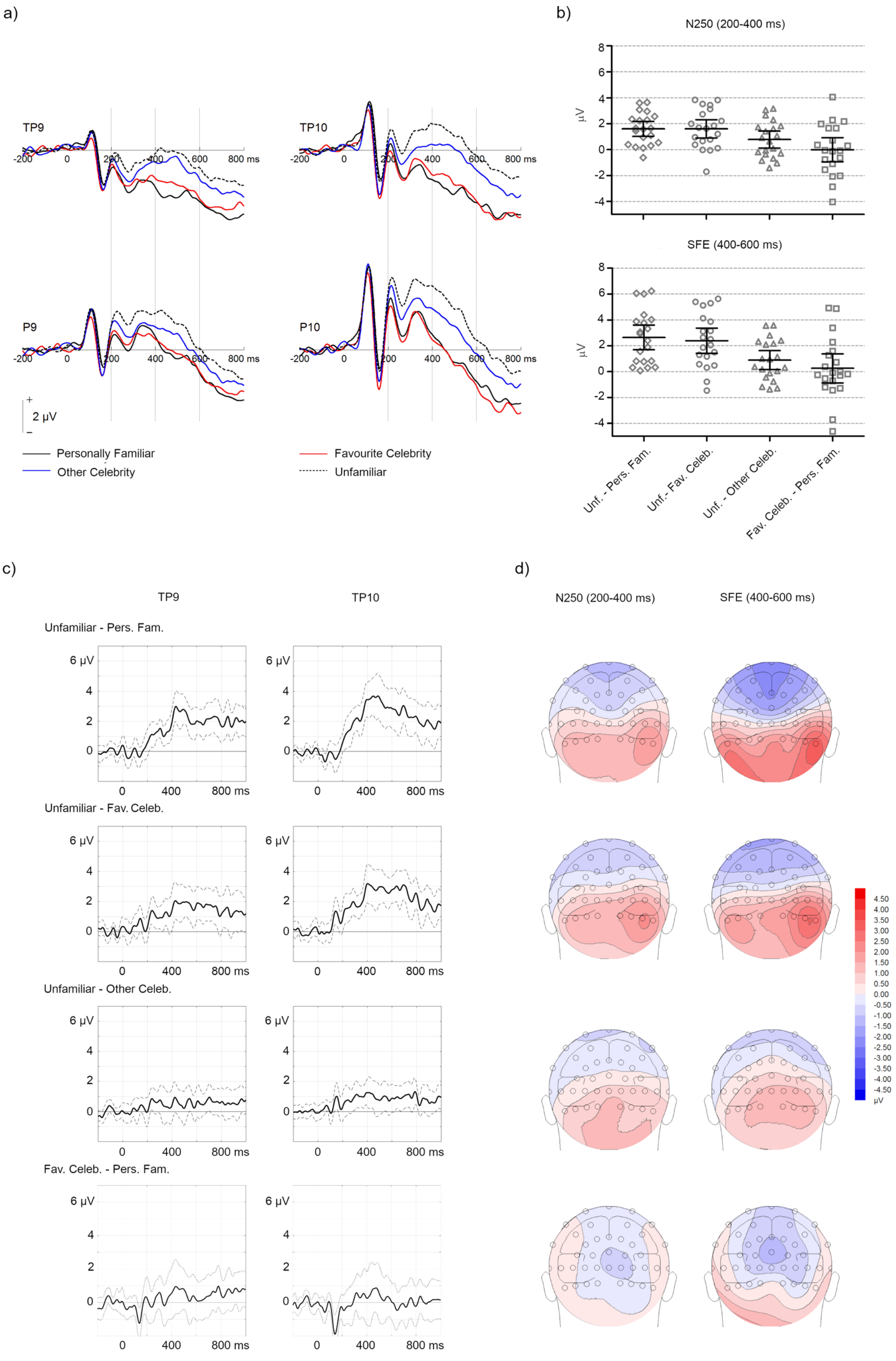

Figure 3. a) Grand average event-related potentials at left- and right-hemispheric occipito-temporal electrodes TP9/TP10 and P9/P10. Dashed lines mark the N250 (200-400 ms) and SFE (400-600 ms) time ranges. b) Mean (+/- 95% CI) and individual familiarity effects in the N250 and SFE time ranges at electrodes TP9/TP10/P9/P10. c) Mean (+/- 95% CI) difference waves at left and right occipito-temporal electrodes TP9/TP10. d) Scalp-topographical voltage maps (spherical spline interpolation, 90° equidistant projection) of familiarity effects in the N250 and SFE time window.

Bootstrapping in the N250 time range revealed reliable effects in 6/20 participants for personally familiar versus unfamiliar faces, P=.40, 95% CI [.22, .61], in 6/20 participants for favourite celebrities versus unfamiliar faces, P=.40, 95% CI [.22, .61], in 4/20 participants for other celebrities versus unfamiliar faces, P=.20, 95% CI [.08, .42], and in 5/20 participants for personally familiar versus favourite celebrity faces, P = .25, 95% CI [.11, .47]. Corresponding analyses in the SFE time range yielded reliable effects in 12/20 participants for personally familiar versus unfamiliar faces, P=.60, 95% CI [.39, .78], in 13/20 participants for favourite celebrities versus unfamiliar faces, P=.65, 95% CI [.43, .82], in 5/20 participants for other celebrities versus unfamiliar faces, P=.25, 95% CI [.11, .47], and in 4/20 participants for personally familiar versus favourite celebrity faces, P = .20, 95% CI [.08, .42].

Rating task

Rating results are reported in Table 1. Personally familiar, favourite celebrities and other celebrity faces were all rated more familiar than unfamiliar faces; personally familiar versus unfamiliar: $t(18) = 8.25$, $p < .001$, $d_{unb.} = 2.56$, favourite celebrities versus unfamiliar: $t(18) = 7.84$, $p < .001$, $d_{unb.} = 2.48$, other celebrity versus unfamiliar: $t(18) = 2.19$, $p = .042$, $d_{unb.} = 0.71$, 95% CI [0.03, 1.45]. Personally familiar and favourite celebrities did not differ significantly, $t(18) = 1.00$, $p = .331$, $d_{unb.} = 0.31$, 95% CI [-0.32, 0.97].

Valence was significantly more positive for personally familiar versus unfamiliar faces, $t(18) = 9.63$, $p < .001$, $d_{unb.} = 3.23$, and favourite celebrities versus unfamiliar faces, $t(18) = 6.87$, $p < .001$, $d_{unb.} = 2.06$. Other celebrity and unfamiliar faces did not differ, $t(18) = 0.59$, $p = .563$, $d_{unb.} = 0.17$, 95% CI [-0.42, 0.77]. Personally familiar faces were rated more positively relative to favourite celebrities, $t(18) = 2.38$, $p = .028$, $d_{unb.} = 0.83$, 95% CI [0.09, 1.62].

Finally, personally familiar, favourite celebrities and other celebrity faces were all rated as more arousing than unfamiliar faces; personally familiar versus unfamiliar: $t(18) = 9.59$, $p < .001$, $d_{unb.} = 2.98$, favourite celebrities versus unfamiliar: $t(18) = 8.81$, $p < .001$, $d_{unb.} = 2.73$, other celebrity versus unfamiliar, $t(18) = 3.15$, $p = .006$, $d_{unb.} = 0.86$, 95% CI [0.26, 1.51]. Personally familiar and favourite celebrities did not differ significantly, $t(18) = 1.30$, $p = .209$, $d_{unb.} = 0.42$, 95% CI [-0.24, 1.10].

Discussion

Experiment 2 replicated the most important finding from Experiment 1 by demonstrating substantial ERP familiarity effects for favourite celebrities. Accordingly, it appears that famous and personally familiar faces are similarly represented, given a sufficient degree of familiarity. In contrast to Experiment 1, however, personally familiar faces did not elicit a larger SFE than favourite celebrities. It should be noted that the increase in effect size found in Experiment 1 was relatively small and therefore presumably difficult to replicate. While further experiments are necessary to unequivocally clarify the exact relationship between the SFE in these different conditions, the primary aim of the present study to obtain further evidence for an SFE elicited by famous faces was clearly met.

Other celebrities in the present experiment also elicited small but significant N250 effects and SFEs, which is different from our previous study (Wiese, Tuttenberg, et al., 2019). As noted above, an interesting difference to this previous experiment is that in Experiment 2 “other” celebrities were not chosen by the experimenters but by the other participant in each pair, and therefore by a person from the same age group and with a more similar educational background. Arguably, celebrities chosen by peers will on average be more familiar than celebrities chosen by the experimenters. At the same time, the SFE for other celebrities was relatively small. It therefore seems reasonable again to conclude that familiar identities elicit a graded SFE[2], with increasing effects elicited by increasingly familiar identities.

Finally, Experiment 2 also found effects involving the familiarity factor in the N170. Both favourite and other celebrities elicited more negative N170s than unfamiliar faces, but personally familiar faces did not. It thus appears that systematic differences between famous and non-famous image sets rather than familiarity per se may underlie this particular result, which is in line with the finding of an "identity type" main effect in the N170 in Experiment 1.

In conclusion, the first two experiments clearly demonstrate that N250 familiarity effects and SFEs can be elicited by famous faces, given that participants are sufficiently familiar with them. However, in both experiments the personally familiar and favourite celebrity faces chosen by participants were ones to which they felt positively towards, as the valence ratings presented in Table 1 confirm. Experiment 3 was therefore designed to explore the relationship between familiarity and affective information as potential characteristics reflected by the SFE.

Experiment 3: Favourite versus disliked celebrities

While both Experiments 1 and 2 show clear SFEs for celebrity faces, it is important to achieve a better understanding of what processes underlie these effects. We have argued before that the SFE presumably reflects the integration of visual with other identity-specific information, such as semantic, episodic or affective information related to an individual person (Wiese, Ingram, et al., 2019; Wiese, Tuttenberg, et al., 2019). We expected that both

[2] Note that even though personally familiar and favourite celebrities were not significantly different in Experiment 2, the pattern of effect sizes was similar to Experiment 1, which is in line with a graded effect in the SFE.

favourite celebrities and personally familiar faces would be highly liked, and that positive valence would be similar in these two conditions. Our rating results in Experients 1 and 2 confirmed this assumption. While any potential difference was therefore unlikely to be explained by valence, it is easily conceivable that positive affect might be a necessary pre-condition to obtain an SFE, or even that the effect is mostly driven by positive affective information rather than familiarity. Experiment 3 therefore attempted to examine the potential role of positive valence for the generation of the effect and to disentangle familiarity and positive affect by introducing an additional disliked celebrity condition.

Based on our previous findings, we expected to find clear ERP familiarity effects for favourite celebrities. The critical question for Experiment 3 was whether similar effects would be observed for a celebrity who is well-known, but disliked. In contrast to Experiments 1 and 2, which kept valence comparable between the critical familiarity conditions, Experiment 3 directly contrasted positive and negative valence while trying to keep familiarity comparable. If the SFE more directly reflects familiarity than affective information, we expected to find a clear SFE for disliked celebrities.

Methods

Participants

Given the somewhat smaller effect size of the SFE for favourite celebrities relative to personally familiar faces in previous experiments, the sample size for Experiment 3 was slightly increased[3]. Twenty-seven Durham University undergraduate students were tested, two of whom were excluded due to technical problems during EEG recording. The final

[3] We did not conduct a formal power analysis before testing Experiment 3. Please note, however, that dz for favourite celebrities versus unfamiliar faces was 1.37 in Experiment 1 and 1.15 in Experiment 2. A post-hoc sensitivity test (repeated-measures t-test, one-sided, power = .95, N = 25) suggests a dz of 0.68, which appears adequate, even when considering that we anticipated the effect to be somewhat smaller. Assuming a less conservative but more conventional power of .8, the experiment had sufficiently large N for an effect of dz = 0.51.

sample consisted of 16 females and nine males, with a mean age of 20.7 years (SD = 0.7). Reimbursement and inclusion/exclusion criteria were identical to the previous experiments. All participants gave written informed consent and the experiment was approved by the ethics committee at Durham University's Psychology department.

Stimuli, procedure, EEG recording and data analysis

For Experiment 3, each participant was asked to name their favourite and least favourite celebrity before the experimental session. Images were collected by the experimenters using Google Images (50 images per ID). Moreover, each individual experiment was completed by adding 50 images of an unfamiliar celebrity and 50 images of a non-famous unfamiliar ID (see Experiment 1). All other aspects of the experiment, including EEG recording and data analysis parameters remained unchanged. Average numbers of analysed trials were 46.0 (SD = 3.2, min = 39) for favourite celebrities, 45.7 (SD = 3.8, min = 36) for disliked celebrities, 46.2 (SD = 4.1, min = 38) for unfamiliar celebrities, and 47.2 (SD = 2.8, min = 40) for unfamiliar faces.

Results

Performance

Performance during the butterfly task was again close to ceiling, mean hit rate = .96, SD = .11, mean false alarm rate < .01, SD = .01. Mean reaction time for correct responses was 545 ms, SD = 71.

Event-related potentials

ERP waveforms showed clear familiarity effects for favourite celebrities and evident but somewhat reduced effects for disliked celebrities (see Figure 4). A repeated-measures ANOVA in the N170 time range did not result in any significant effects involving the familiarity factor, all $F < 1.93$, all $p > .133$, all $\eta^2_p < .074$. A corresponding analysis in the

N250 time range yielded a trend for a significant main effect of familiarity, $F(3, 72) = 2.21$, $p = .094$, $\eta^2_p = .084$, 90% CI [.0, .169]. Planned comparisons revealed significantly more negative amplitudes for the favourite celebrities versus unfamiliar condition, $M_{diff.} = 0.72$ µV, 95% CI [0.21, 1.24], $t(24) = 2.93$, $p = .007$, $d_{unb.} = 0.22$, 95% CI [0.06, 0.39], and for the unfamiliar celebrity versus unfamiliar condition, $M_{diff.} = 0.63$ µV, 95% CI [0.03, 1.23], $t(24) = 2.15$, $p = .042$, $d_{unb.} = 0.19$, 95% CI [0.01, 0.37]. Disliked celebrities did not differ significantly from the unfamiliar, $M_{diff.} = 0.53$µV, 95% CI [-0.19, 1.25], $t(24) = 1.51$, $p = .144$, $d_{unb.} = 0.15$, 95% CI [-0.05, 0.35], or favourite celebrities condition, $M_{diff.} = 0.20$ µV, 95% CI [-0.33, 0.73], $t(24) = 0.77$, $p = .449$, $d_{unb.} = 0.05$, 95% CI [-0.09, 0.19]. An ANOVA using vector-length based corrected N250 amplitudes did not detect any significant interaction of familiarity with the hemisphere or site factors, all $F < 2.08$, all $p > .110$, all $\eta^2_p < .081$.

A corresponding analysis in the SFE time window revealed a significant main effect of familiarity, $F(3, 72) = 9.65$, $p < .001$, $\eta^2_p = .287$, 90% CI [.126, .392]. The SFE was more negative for favourite celebrities versus unfamiliar, $M_{diff.} = 1.67$ µV, 95% CI [0.97, 2.38], $t(24) = 4.90$, $p < .001$, $d_{unb.} = 0.53$, 95% CI [0.28, 0.81], and for disliked celebrities versus unfamiliar faces, $M_{diff.} = 0.96$ µV, 95% CI [0.15, 1.77], $t(24) = 2.45$, $p = .022$, $d_{unb.} = 0.30$, 95% CI [0.04, 0.57]. Unfamiliar celebrities did not differ significantly from unfamiliar faces, $M_{diff.} = 0.21$ µV, 95% CI [-0.39, 0.81], $t(24) = 0.72$, $p = .476$, $d_{unb.} = 0.08$, 95% CI [-0.14, 0.30]. Favourite celebrities elicited significantly more negative amplitudes relative to disliked celebrities, $M_{diff.} = 0.72$ µV, 95% CI [0.14, 1.29], $t(24) = 2.57$, $p = .017$, $d_{unb.} = 0.20$, 95% CI [0.04, 0.37]. Again, an ANOVA using vector-length corrected SFE amplitudes did not detect any significant interaction of familiarity with the hemisphere or site factors, all $F < 1.93$, all $p > .134$, all $\eta^2_p < .074$.

A further ANOVA on N250-corrected SFE measures yielded a significant main effect of familiarity, $F(3, 72) = 17.39$, $p < .001$, $\eta^2_p = .420$, 90% CI [.256, .516]. Corrected SFE was more negative for favourite celebrities versus unfamiliar faces, $M_{diff.} = 0.95$ µV, 95% CI [0.49, 1.41], $t(24) = 4.24$, $p < .001$, $d_{unb.} = 0.54$, 95% CI [0.25, 0.86]. Disliked celebrities were more negative than non-famous unfamiliar faces, $M_{diff.} = 0.43$ µV, 95% CI [0.11, 0.75], $t(24) = 2.81$, $p = .010$, $d_{unb.} = 0.23$, 95% CI [0.06, 0.42], and unfamiliar celebrities, $M_{diff.} = 0.85$ µV, 95% CI [0.47, 1.23], $t(24) = 4.61$, $p < .001$, $d_{unb.} = 0.48$, 95% CI [0.24, 0.75]. Unfamiliar celebrities were significantly less negative than unfamiliar faces, $M_{diff.} = -0.42$ µV, 95% CI [-0.71, -0.13], $t(24) = -3.01$, $p = .006$, $d_{unb.} = -0.24$, 95% CI [-0.43, -0.07]. Finally, favourite celebrities elicited significantly more negative amplitudes relative to disliked celebrities, $M_{diff.} = 0.52$ µV, 95% CI [0.10, 0.93], $t(24) = 2.56$, $p = .017$, $d_{unb.} = 0.29$, 95% CI [0.05, 0.54].

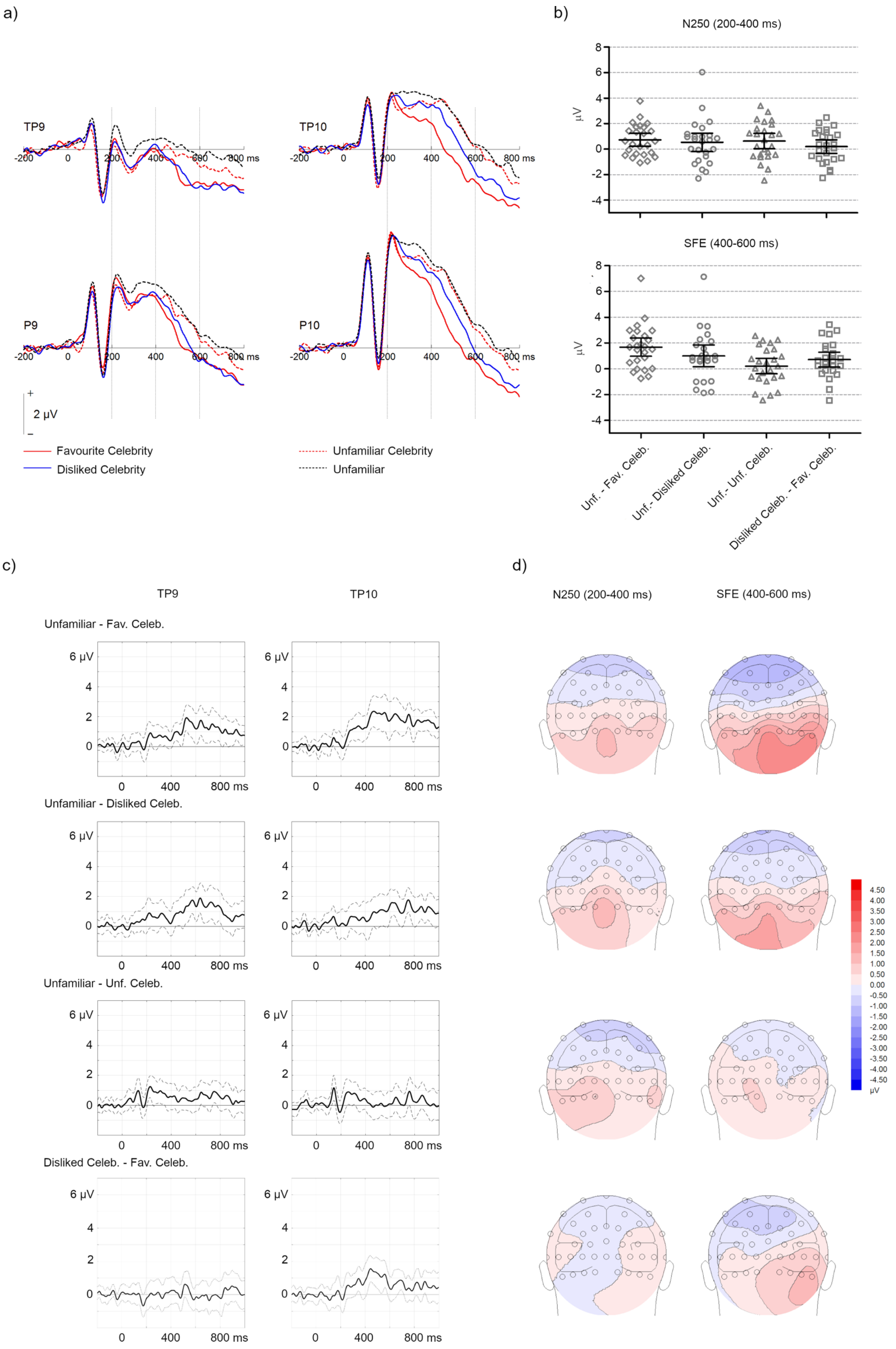

Figure 4. a) Grand average event-related potentials at left- and right-hemispheric occipito-temporal electrodes TP9/TP10 and P9/P10. Dashed lines mark the N250 (200-400 ms) and SFE (400-600 ms) time ranges. b) Mean (+/- 95% CI) and individual familiarity effects in the N250 and SFE time ranges at electrodes TP9/TP10/P9/P10. c) Mean (+/- 95% CI) difference waves at left and right occipito-temporal electrodes TP9/TP10. d) Scalp-topographical voltage maps (spherical spline interpolation, 90° equidistant projection) of familiarity effects in the N250 and SFE time window.

Bootstrapping analyses in the N250 time window revealed reliable effects in 7/25 participants for favourite celebrities versus unfamiliar faces, P = .28, 95% CI [.14, .48], in 7/25 participants for disliked celebrities versus unfamiliar faces, P = .28, 95% CI [.14, .48], in 6/25 participants for unfamiliar celebrities versus unfamiliar faces, P = .24, 95% CI [.12, .43], and in 2/25 participants for favourite versus disliked celebrity faces, P = .08, 95% CI [.02, .25]. Corresponding analyses in the SFE time range yielded reliable effects in 11/25 participants for favourite celebrities versus unfamiliar faces, P = .44, 95% CI [.27, .63], in 8/25 for disliked celebrities versus unfamiliar faces, P = .32, 95% CI [.17, .52], in 5/25 for unfamiliar celebrities versus unfamiliar faces, P = .20, 95% CI [.09, .39], and in 4/20 participants for favourite versus disliked celebrity faces, P = .16, 95% CI [.06, .35].

Rating task

Rating results are reported in Table 1. Favourite celebrities were rated as more familiar than unfamiliar faces, $t(24) = 12.63$, $p < .001$, $d_{unb.} = 3.40$, which was also the case for disliked celebrities versus unfamiliar faces, $t(24) = 10.80$, $p < .001$, $d_{unb.} = 2.86$. Unfamiliar celebrities were not rated as more familiar than unfamiliar faces, $t(24) = 1.70$, $p = .102$, $d_{unb.} = 0.44$, 95% CI [-0.09, 0.98], nor did favourite celebrities differ from disliked celebrities, $t(24) = 1.55$, $p = .134$, $d_{unb.} = 0.37$, 95% CI [-0.12, 0.88].

As expected, favourite celebrities were rated significantly more positive relative to unfamiliar faces, $t(24) = 7.49$, $p < .001$, $d_{unb.} = 2.17$, and disliked celebrities were rated as more negative, $t(24) = 6.73$, $p < .001$, $d_{unb.} = 1.92$, 95% CI [1.17, 2.77]. Unfamiliar celebrities and unfamiliar faces did not differ, $t(24) = 0.81$, $p = .425$, $d_{unb.} = 0.22$, 95% CI [-0.33, 0.78]. Moreover, favourite celebrities were rated as significantly more positive than disliked celebrities, $t(24) = 12.84$, $p < .001$, $d_{unb.} = 4.29$.

Finally, favourite celebrities were rated as more arousing than unfamiliar faces, $t(24) = 8.05$, $p < .001$, $d_{unb.} = 1.88$, 95% CI [1.21, 2.66], which was also observed for disliked celebrities, $t(24) = 5.33$, $p < .001$, $d_{unb.} = 1.53$, 95% CI [0.84, 2.30]. Unfamiliar celebrities did not differ from unfamiliar faces, $t(24) = 1.10$, $p = .282$, $d_{unb.} = 0.21$, 95% CI [-0.18, 0.62], nor did favourite celebrities differ from disliked celebrities, $t(24) = 1.00$, $p = .327$, $d_{unb.} = 0.24$, 95% CI [-0.24, 0.73].

Discussion

With an overall aim of beginning to disentangle positive valence and familiarity, Experiment 3 tested whether well-known but disliked celebrities would elicit similar ERP familiarity effects to favourite celebrities. The results showed a significant SFE for disliked celebrities, and it thus appears that positive valence is not a necessary pre-requisite to elicit the effect. However, the SFE for disliked celebrities was reduced relative to the corresponding effect for favourite celebrities. This smaller effect may be related to less identity-specific information available for disliked celebrities. Arguably, even though disliked people are familiar, participants likely know more about the celebrities they particularly like, as they may actively seek information about them while at the same time avoiding disliked people. Notably, however, the less liked identities in Experiment 3 elicited the smaller SFEs, which might still indicate that the magnitude of the effect partly reflects positive valence. For instance, it remains possible that a part of the effect is driven by familiarity, but that it is boosted by positive affective information. We will return to this point in Experiment 4.

Experiment 3 did not show a significant N250 familiarity effect for disliked celebrities. Although the direction and timing of such effects is very consistent across experiments (see Figures 1-3; Wiese, Ingram, et al., 2019; Wiese, Tuttenberg, et al., 2019), they are generally small for (non-favourite) celebrity faces. Statistical significance for these N250 effects will therefore be observed in some, but not all experiments and conditions, given typical sample sizes for ERP research. Moreover, an N250 effect was observed for unfamiliar celebrities, and we have discussed potential explanations for such "false familiarity effects" in Experiment 1. Together these findings suggest that, relative to other ERP markers such as the SFE or the N250r (e.g., Schweinberger & Neumann, 2016; Schweinberger, Pickering, Jentzsch, Burton, & Kaufmann, 2002), N250 familiarity effects are more variable and presumably influenced by factors additional to familiarity per se.

Experiment 4: Own versus personally familiar faces

Overall, the findings of Experiments 1-3 suggest that celebrity faces elicit qualitatively similar ERP familiarity effects relative to personally familiar faces, suggesting that media-based and personal familiarity rely on the same type of representation. Experiment 4 sought to extend this suggestion further to the idea that *all* familiar faces are represented in a qualitatively similar way. This was achieved by examining ERP correlates of an additional potential type of familiarity, namely the observer's own face. Recognition of one's own face occurs in various everyday contexts (Bredart & Young, 2004), and has been linked to the development of a self-concept (Devue & Bredart, 2011; Gallup, 1970; Keenan, Wheeler, Gallup, & Pascual-Leone, 2000). In line with the idea that there is something special about self-recognition, previous ERP studies have suggested that own face processing is accompanied by distinct neural processes (Butler, Mattingley, Cunnington, & Suddendorf, 2013; Keyes, Brady, Reilly, & Foxe, 2010). This interpretation seems to be interestingly different from the pattern emerging from the experiments reported here, which have not yet found any evidence of different "types" of familiarity. Therefore, using the approach of testing participants with multiple ambient images per identity, Experiment 4 aimed at testing whether own and personally familiar faces are represented in a qualitatively similar way.

In addition, examining own face recognition offers a different perspective on previous ideas about the processes underlying the SFE. We have previously suggested that the SFE might be related to the preparation for an interaction (Wiese, Ingram, et al., 2019). While Experiments 1-3 seem to contradict this idea, as we do not interact with people exclusively known via media, one could argue that our face recognition system has evolved in an environment in which it did not have to distinguish between real-life and media-based familiarity. It might therefore be possible that any process involved in preparing for a potential interaction is automatically triggered whenever a face of a different person is perceived. Recognition of a face as one's own, however, should not trigger this process.

Experiment 4 also revisits the question of whether the SFE is related to positive valence. Experiment 3 revealed a smaller SFE for disliked relative to favourite celebrities, and we interpreted this finding as reflecting reduced availability of identity-specific information in the former condition. This experiment alone, however, was not able to completely disentangle positive affect and familiarity, as the less positive face was also likely the lesser known. For Experiment 4, we reasoned that the own face is arguably the most familiar face possible. At the same time, participants in Experiments 1-3 sometimes mentioned that they do not particularly like seeing pictures of themselves. Accordingly, and in contrast to Experiment 3, the less positive face could be *more* well-known when participants were tested with own and personally familiar faces.

We therefore predicted that the SFE for own faces should be smaller relative to the effect observed for personally familiar faces if the effect is (partly) driven by either positive valence or the preparation for an interaction. If, however, the SFE is mostly driven by familiarity and the integration of person-specific knowledge, the effect should be larger for own faces.

Methods

Participants

Twenty-two undergraduate students at Durham University were tested, two of which were excluded due to technical problems during EEG recording. The final sample consisted of 17 females and three males, with a mean age of 19.9 years (SD = 1.5). Reimbursement and inclusion/exclusion criteria were identical to the previous experiments. All participants gave written informed consent and the experiment was approved by the ethics committee at Durham University's Psychology department.

Stimuli, procedure, EEG recording and data analysis

For Experiment 4, each participant provided 50 images of a highly personally familiar ID (not known from university) and 50 images of their own face. Participants were paired, and personally familiar and own face IDs for one participant in each pair were used as the unfamiliar and "other face" IDs, respectively, for the other participant. Accordingly, while personally familiar faces of one participant in a given pair were unfamiliar to the other participant, this was not controlled for the own face condition. In other words, the own face of one participant in a given pair may have been familiar to the other participant, given that all participants were students at Durham University. Note, however, that familiarity with all IDs was rated after the main experiment. All other aspects of the experiment, including EEG recording and data analysis parameters remained unchanged. Average numbers of analysed trials were 47.5 (SD = 3.5, min = 39) for own faces, 47.5 (SD = 3.2, min = 40) for personally familiar faces, 47.7 (SD = 2.6, min = 40) for other faces, and 47.5 (SD = 2.9, min = 40) for unfamiliar faces.

Results

Performance

Performance during the butterfly task was again close to ceiling, mean hit rate = .98, SD = .03, mean false alarm rate < .01, SD = .002. Mean reaction time for correct responses was 555 ms, SD = 93.

Event-related potentials

ERP waveforms revealed very large familiarity effects for the participants' own face, as well as relatively smaller, but still large effects for personally familiar faces (see Figure 5). A repeated-measures ANOVA in the N170 time range did not yield any significant effects involving the familiarity factor, all $F < 2.63$, all $p > .059$, all $\eta^2_p < .122$. Analysis of the N250 time range (200-400 ms) revealed a significant main effect of familiarity, $F(3, 57) = 67.30$, $p < .001$, $\eta^2_p = .780$, 90% CI [.679, .823], as well as an interaction of familiarity by hemisphere, $F(3, 57) = 17.56$, $p < .001$, $\eta^2_p = .480$, 90% CI [.297, .576]. Familiarity effects were larger over the right hemisphere for the participants' own, $F(1, 19) = 31.02$, $p < .001$, $\eta_p^2 = .620$, 90% CI [.342, .740], and for personally familiar faces, $F(1, 19) = 5.00$, $p = .037$, $\eta_p^2 = .208$, 90% CI [.007, .427]. Moreover, planned comparisons showed that the N250 was significantly more negative for personally familiar versus unfamiliar faces, $M_{diff.} = 2.14$ µV, 95% CI [1.50, 2.78], $t(19) = 6.99$, $p < .001$, $d_{unb.} = 0.53$, 95% CI [0.32, 0.79], as well as own versus unfamiliar faces, $M_{diff.} = 3.89$ µV, 95% CI [3.14, 4.64], $t(19) = 10.87$, $p < .001$, $d_{unb.} = 1.04$, 95% CI [0.68, 1.47]. Other faces (i.e., the "own face" of a different participant) and unfamiliar faces did not differ significantly, $M_{diff.} = 0.08$ µV, 95% CI [-0.41, 0.57], $t(19) = 0.35$, $p = .730$, $d_{unb.} = 0.02$, 95% CI [-0.10, 0.15]. The own-face condition was significantly more negative than the personally familiar condition, $M_{diff.} = 1.75$ µV, 95% CI [1.24, 2.25], $t(19) = 7.24$, $p < .001$, $d_{unb.} = 0.47$, 95% CI [0.28, 0.69]. A further ANOVA using vector-length corrected N250 amplitudes yielded a significant interaction of familiarity by hemisphere, $F(3, 57) = 7.09$, $p < .001$, $\eta^2_p = .272$, 90% CI [.093, .387]. However, no significant hemisphere effects were observed for any of the four experimental conditions when tested separately, all $F < 2.06$, all $p > .167$, all $\eta^2_p < .099$.

A corresponding ANOVA in the SFE time window again yielded a significant main effect of familiarity, $F(3, 57) = 92.40$, $p < .001$, $\eta^2_p = .829$, 90% CI [.749, .863], as well as significant interactions of site by familiarity, $F(3, 57) = 3.72$, $p = .016$, $\eta^2_p = .164$, 90% CI [.018, .274], and hemisphere by familiarity, $F(3, 57) = 22.39$, $p < .001$, $\eta^2_p = .541$, 90% CI [.367, .627]. Familiarity effects were larger over the right hemisphere for both the participants' own, $F(1, 19) = 37.27$, $p < .001$, $\eta_p^2 = .662$, 90% CI [.400, .769], and for personally familiar faces, $F(1, 19) = 17.67$, $p < .001$, $\eta_p^2 = .482$, 90% CI [.184, .643]. Planned comparisons revealed significantly more negative amplitudes for personally familiar relative to unfamiliar faces, $M_{diff.} = 3.30$ µV, 95% CI [2.31, 4.29], $t(19) = 6.95$, $p < .001$, $d_{unb.} = 0.87$, 95% CI [0.52, 1.29], as well as for own-face versus unfamiliar faces, $M_{diff.} = 6.32$ µV, 95% CI [5.20, 7.49], $t(19) = 11.77$, $p < .001$, $d_{unb.} = 1.59$, 95% CI [1.06, 2.24]. Again, the other face condition did not differ significantly from the unfamiliar condition, $M_{diff.} = 0.20$ µV, 95% CI [-0.37, 0.77], $t(19) = 0.72$, $p = .478$, $d_{unb.} = 0.05$, 95% CI [-0.10, 0.21], while the own-face condition was more negative than the personally familiar condition, $M_{diff.} = 3.02$ µV, 95% CI [2.50, 3.55], $t(19) = 12.15$, $p < .001$, $d_{unb.} = 0.77$, 95% CI [0.52, 1.09]. An ANOVA using vector-length corrected SFE amplitudes again revealed a significant interaction of familiarity by hemisphere, $F(3, 57) = 5.02$, $p = .004$, $\eta^2_p = .209$, 90% CI [.046, .323]. However, again no significant hemisphere effects were observed for any of the four experimental conditions when tested separately, all $F < 2.06$, all $p > .167$, all $\eta^2_p < .099$.

A further ANOVA on N250-corrected SFE again yielded a significant main effect of familiarity, $F(3, 57) = 34.44$, $p < .001$, $\eta^2_p = .644$, 90% CI [.496, .713], as well as significant interaction of site by familiarity, $F(3, 57) = 14.75$, $p < .001$, $\eta^2_p = .437$, 90% CI [.249, .538], and hemisphere by familiarity, $F(3, 57) = 4.07$, $p = .011$, $\eta^2_p = .176$, 90% CI [.025, .289]. Planned comparisons revealed significantly more negative amplitudes for personally familiar relative to unfamiliar faces, $M_{diff.} = 1.16$ µV, 95% CI [0.56, 1.76], $t(19) = 4.05$, $p = .001$, $d_{unb.} = 0.62$, 95% CI [0.27, 1.02], as well as for own-face versus unfamiliar faces, $M_{diff.} = 2.44$ µV, 95% CI [1.73, 3.14], $t(19) = 7.23$, $p < .001$, $d_{unb.} = 1.19$, 95% CI [0.72, 1.75]. Again, the other face condition did not differ significantly from the unfamiliar condition, $M_{diff.} = 0.11$ µV, 95% CI [-0.16, 0.39], $t(19) = 0.87$, $p = .397$, $d_{unb.} = 0.06$, 95% CI [-0.08, 0.21], but the own-face condition was more negative than the personally familiar condition, $M_{diff.} = 1.28$ µV, 95% CI [0.75, 1.81], $t(19) = 5.05$, $p < .001$, $d_{unb.} = 0.60$, 95% CI [0.31, 0.94].

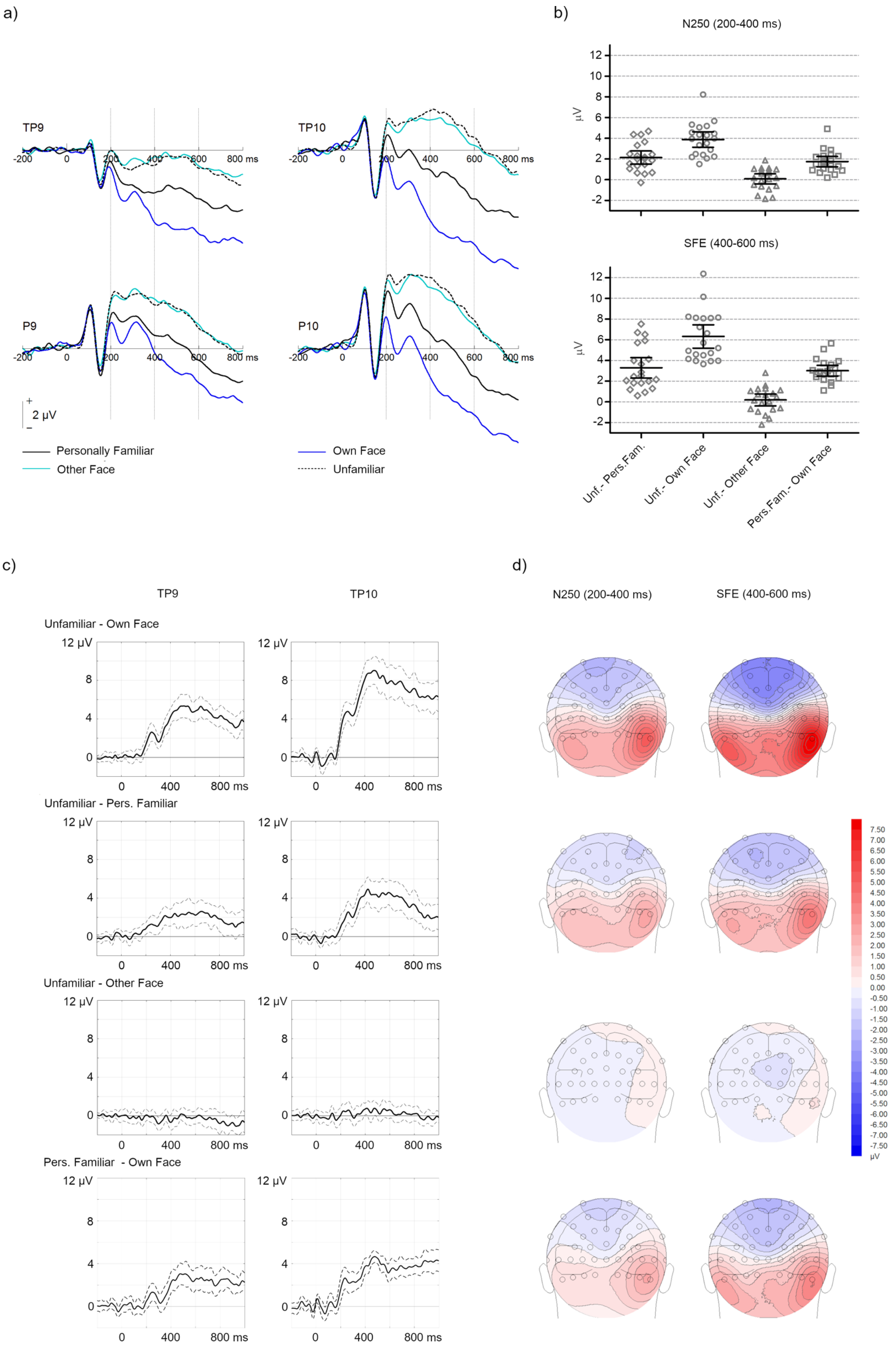

Figure 5. a) Grand average event-related potentials at left- and right-hemispheric occipito-temporal electrodes TP9/TP10 and P9/P10. Dashed lines mark the N250 (200-400 ms) and SFE (400-600 ms) time ranges. b) Mean (+/- 95% CI) and individual familiarity effects in the N250 and SFE time ranges at electrodes TP9/TP10/P9/P10. c) Mean (+/- 95% CI) difference waves at left and right occipito-temporal electrodes TP9/TP10. d) Scalp-topographical voltage maps (spherical spline interpolation, 90° equidistant projection) of familiarity effects in the N250 and SFE time window.

Bootstrapping analysis in the N250 time range revealed reliable familiarity effects in 20/20 participants for the participants' own versus unfamiliar faces, P = 1.0, 95% CI [.84, 1.0], in 13/20 participants for personally familiar versus unfamiliar faces, P = .65, 95% CI [.43, .82], in 1/20 participants for other relative to unfamiliar faces, P = .05, 95% CI [.01, .24], and in 10/20 participants for own versus personally familiar faces, P = .50, 95% CI [.30, .70]. A corresponding analysis in the SFE time window yielded reliable familiarity effects in 20/20 participants for the participants' own versus unfamiliar faces, P = 1.0, 95% CI [.84, 1.0], in 14/20 participants for personally familiar versus unfamiliar faces, P = .70, 95% CI [.48, .86], in 1/20 participants for other relative to unfamiliar faces, P = .05, 95% CI [.01, .24], and in 18/20 participants for own versus personally familiar faces, P = .90, 95% CI [.70, .97].

Rating task

Results from the rating task are reported in Table 1. Participants rated both their own, $t(19) = 12.35$, $p < .001$, $d_{unb} = 3.75$, and the personally familiar face, $t(19) = 11.71$, $p < .001$, $d_{unb} = 3.42$, as more familiar relative to the unfamiliar face. At the same time, rated familiarity did not differ for the other versus unfamiliar face, $t(19) = 1.99$, $p = .061$, $d_{unb} = 0.36$, 95% CI [-0.02, 0.76], nor for the own versus personally familiar face, $t(19) = 1.00$, $p = .330$, $d_{unb} = 0.30$, 95% CI [-0.32, 0.94].

Importantly, personally familiar faces were rated more positively both relative to the own face, $t(19) = 3.90$, $p = .001$, $d_{unb} = 0.65$, 95% CI [0.27, 1.08], and the unfamiliar face, $t(19) = 3.94$, $p = .001$, $d_{unb} = 1.47$, 95% CI [0.61, 2.41]. Neither the own face, $t(19) = 1.81$, $p = .086$, $d_{unb} = 0.65$, 95% CI [-0.02, 0.76], nor the other face, $t(19) = -0.30$, $p = .772$, $d_{unb} = -0.06$, 95% CI [-0.51, 0.38], was rated differently relative to the unfamiliar face.

Both own, $t(19) = 3.16$, $p = .005$, $d_{unb} = 1.08$, 95% CI [0.33, 1.90], and personally familiar faces, $t(19) = 5.21$, $p < .001$, $d_{unb} = 1.79$, 95% CI [0.94, 2.77], were rated as more arousing than unfamiliar faces. Moreover, personally familiar faces were rated as more arousing than own faces, $t(19) = 2.77$, $p = .012$, $d_{unb} = 0.63$, 95% CI [0.14, 1.16]. Finally, other and unfamiliar faces did not differ, $t(19) = 1.42$, $p = .171$, $d_{unb} = 0.20$, 95% CI [-0.09, 0.50].

Discussion

Experiment 4 was conducted to test further hypotheses about the nature of familiar face representations and the processes underlying the SFE. Large N250 familiarity effects and SFEs were found for own relative to unfamiliar faces. Although different in amplitude, these ERP familiarity effects are highly similar with respect to timing and scalp-distribution relative to the corresponding effects for personally familiar faces. We therefore suggest that own faces are not represented differently and elicit qualitatively similar neural processes relative to personally familiar faces. Having now tested three potential “types” of familiarity, it appears as if familiarity varies in strength for different identities, but relies on the same neural mechanisms.

ERP familiarity effects were more pronounced for own relative to personally familiar faces (see also Butler et al., 2013; Keyes et al., 2010), and both effect sizes in the group analysis and reliabilities in the bootstrapping analyses for individual subjects were impressive in the former condition. Importantly, the larger SFE for own relative to personally familiar faces does not sit easily with potential explanations in terms of enhanced affective processing, as own faces were rated as being both less arousing and less positive than personally familiar faces. This latter finding is particularly noteworthy in combination with the results of Experiment 3, which found that disliked celebrities elicited smaller SFEs relative to favourite celebrities. The very large SFE for own faces in Experiment 4 also strongly argues against the suggestion that the effect reflects the preparation of an interaction. We will elaborate on these points in the general discussion.

General Discussion

To better understand mental representations underlying face familiarity, the present experiments used ERPs as a sensitive measure of time-locked neural responses to a variety of different face categories, varying in their degree and type of familiarity (real-life versus media-based), as well as their emotional valence. The N250 familiarity effect was used as an index of visual familiarity and the recently discovered SFE as an index involving integration of identity-specific information. Our approach offered a strong test of recognition by using multiple highly variable everyday ambient images (Burton et al., 2016; Kramer et al., 2018). This was done because the observation of a high degree of image invariance, in the sense that almost any image of a highly familiar face will be recognised with ease, is central to understanding familiar face recognition (Bruce & Young, 1986; Burton et al., 2016). The main findings clearly reveal degree rather than type of familiarity as the major principle underlying responses to face familiarity. They further show that familiar face representations become activated at both ends of the continuum of perceived valence, i.e., whether the person is liked or disliked.

Experiments 1 and 2 measured the N250 and SFE to personally familiar faces and favourite celebrities in comparison to different types of unfamiliar faces. Stimuli were tailored to individual participants throughout these experiments, allowing a level of analysis which is not possible in studies which manipulate familiarity as a simple binary (familiar vs. unfamiliar) contrast. Across both experiments a clear N250 familiarity effect and SFE was evident to celebrity as well as personally familiar faces, which suggests that *type* of familiarity is not a critical aspect for familiar face representations. Instead, these findings show that robust visual representations of familiar faces can be established even if the specific person is not known from real life but only via media exposure. It thus appears that real-life 3d visual exposure and the possibility to actively explore the visual environment through one's own movements are not necessary for establishing such representations. More generally, the findings also imply that the SFE does not seem to reflect preparation for social interaction, as we do not interact with the celebrities we recognise in photos. Instead, they suggest a key role for bringing to mind pertinent facts and episodes from past experience that facilitate the interpretation of someone's behaviour, which is consistent with a strong influence of *degree* of familiarity.

While Experiments 1 and 2 controlled valence by choosing familiar faces that were all of liked individuals, Experiment 3 then manipulated this key dimension of emotion by using faces of liked (positive valence) or disliked (negative valence) celebrities. There was evidence of an SFE to both liked and disliked celebrities, though this was stronger for the liked celebrities. That the SFE was evident at both ends of the valence dimension suggests that this is not a critical contributor, and the *quantitative* effect may instead reflect people’s tendency to spend more time learning about liked than disliked people.

Experiment 4 then demonstrated a substantially larger SFE for the participants’ own, relative to a personally familiar face. As viewing the own face images was rated as less positive than viewing personally familiar faces, this finding again argues against any straightforward modulation of the SFE by valence. Moreover, an increased SFE for own faces is again not in line with the suggestion that the effect reflects the preparation of a potential interaction. Instead, these findings converge in suggesting a key role for the *degree* of familiarity during the integration of visual and additional identity-specific information.

Taken together, then, our findings demonstrate remarkable consistency in the processing of face familiarity as indexed by the N250 and especially the SFE. These effects are unmoderated by the source of familiarity (personal acquaintance or mass media) and at least partly independent of valence information (i.e., whether we like or dislike a particular person). These conclusions are supported by the consistent finding of an SFE for all well-known facial identities, whether they are celebrities or known from real life, particularly liked or disliked, and whether they depict the participants' own versus another face. The SFE in the present experiments was largest for the own face, followed by personally familiar, favourite celebrity, and disliked celebrity faces. However, both timing and scalp distribution of the SFE for different categories of familiar faces were remarkably similar. This overall result pattern of varying amplitudes but similar scalp distribution and timing seems to reflect the degree of participants' familiarity with these face categories and consequential amounts of identity-specific knowledge.

Although the scalp distributions of the N250 effect and the SFE are highly similar, it seems reasonable to assume that the latter effect does not merely reflect visual familiarity (which should be resolved at the N250 stage; Schweinberger & Neumann, 2016). Instead, the SFE appears to be modulated by the amount of available semantic and/or episodic information (for a review discussing the integration of such information, see Gobbini & Haxby, 2007). We know more about ourselves and highly personally familiar people than famous people, even if they are our favourite celebrities. Similarly, we actively seek information about our favourite celebrities, whereas most of us probably spend less time watching or listening to celebrities we dislike. On the basis of the present and previous findings, it thus appears that the SFE indexes the integration of visual with identity-specific semantic and/or episodic information.

We have suggested previously (Wiese, Ingram, et al., 2019) that the SFE reflects the integration of person-related information needed to prepare for an interaction. The present findings do not seem to be fully in line with this suggestion, since we do not usually interact with people we do not meet in person. However, it may nonetheless be the case that the SFE reflects access to identity-specific information needed to interpret someone's behaviour (such as a politician, or a character in a film), or simply to make sense of a particular ambient image (e.g. "why did I look so grumpy in that photo?").

We further note that the N250 effect was small and did not reach statistical significance for disliked celebrities in Experiment 3. As participants presumably avoid disliked identities to some extent, while actively seeking exposure to highly liked faces, a relatively high degree of familiarity seems necessary to reliably elicit the N250 effect from ambient images. Previous work suggests that the robustness of familiar face representations, i.e., the probability of activating them with a wide range of images, depends on our experience with within-person variability (Burton et al., 2016; Kramer et al., 2018). If this experience is limited for a given identity, the face will not be recognised from all of the presented images. Such partial recognition failures from less “typical” images may explain the small and non-significant N250 effects for lesser known celebrities in our present and previous studies (Wiese, Tuttenberg, et al., 2019). It should also be noted that an N250 effect for unfamiliar celebrities was found in Experiment 3 but not in Experiment 1. This small (with Cohen’s d < 0.2) and inconsistent effect might reflect differences in the nature of pictures of celebrities relative to non-famous unfamiliar faces. For example, celebrities' faces may on average differ from other faces in attractiveness or distinctiveness and they are more likely to be photographed in certain ways. Such differences might lead to slightly higher interest and increased face learning during the experiment. However, the use of comparisons between images of unfamiliar (i.e. unrecognised) celebrities and other unfamiliar faces in most experiments shows that such influences are minor at best.

It is important to consider whether systematic differences between image sets, in addition to familiarity, could have affected our results. Ideally, the same stimuli should be used in all experimental conditions, but this is clearly not possible in the present study (or, indeed, many others in the field). However, we propose that stimulus effects are not systematically detectable in our data, for the following reasons. First, independent of *potential* confounds, our results are *de facto* not affected by differences between the image sets per se. For instance, low- and mid-level visual differences (e.g., in luminance, contrast, spatial frequency spectrum etc.) are visible in early ERP components. In the difference waves of our experiments (see figures 2-5c), the lower boundaries of the 95% confidence intervals were generally below zero until approximately 200 ms after stimulus onset, suggesting no reliable differences during processing stages involved in basic visual processes. Most notably, the P1 component, which peaks approximately between 80 and 120 ms, is widely known to be highly sensitive to low-level visual characteristics. Our results, however, show no systematic differences in this time range that could have substantially affected our results.

Second, this de facto absence of image set confounds is presumably the consequence of our experimental set-up. Here, we have used sets of 50 “ambient” images for each presented identity. As noted above, these images vary “naturally” on a large number of dimensions, which directly affect the low- and mid-level image characteristics. Because variability within each image set is high, systematic differences between sets used in the different experimental conditions are unlikely.

Third, across the experiments, we have used two different strategies to prevent potential image differences by (i) matching face identities with respect to age, gender, hair style and hair colour (Experiments 1 and 3), and (ii) by balancing face IDs across conditions

(Experiments 2 and 4). Using two different strategies appears useful to us, as no single experimental approach can ever be perfect. Converging results coming from two different approaches, however, appear more convincing, and in this case the results are quite consistent.

Finally, we would like to briefly consider the alternative to our approach. In the experiments presented here, we used different identities/image sets for different participants in the respective experimental conditions. If, alternatively, all participants saw the same familiar and unfamiliar faces, it seems plausible that differences between items unrelated to familiarity were *more* likely to drive potential ERP effects relative to our approach. No control for stimulus characteristics can possibly be perfect, and two stimuli that differ in *no* visual feature are arguably identical pictures. Any residual difference between image sets will then be systematically presented to all participants. Even more importantly, heavily controlled stimuli quickly become ecologically invalid, which raises the question whether the cognitive and neural processes engaged in recognising these stimuli are the same as in real-life face recognition (see e.g. Burton, 2013). In comparison, the use of multiple "ambient" images appears clearly preferable.

A further more general point is related to the distinction between quantitative and qualitative differences. For face recognition, most researchers presumably agree that there is a qualitative difference between familiar and completely unfamiliar faces, in the sense that robust face and identity-specific representations can only possibly exist for the former category (Young & Burton, 2017, 2018). Accordingly, given adequate experimental control, a difference between a specific familiar face category and unfamiliar faces is likely to reflect a qualitative difference (i.e., the activation of the familiar face's representation). The starting point for the present series of experiments was the absence of such an effect in our previous study (Wiese, Tuttenberg, et al., 2019) in which we did not find a difference between famous and unfamiliar faces in the SFE. Of course, the inability to measure a response does not necessarily reflect its absence. However, failing to observe evidence for a difference between unfamiliar and famous faces, while at the same time observing substantial familiarity effects for personally familiar faces, would have suggested that the underlying representation is only available for the faces of personally familiar people. This in turn would have been interpreted as a qualitative difference. However, in the experiments presented here, various different categories of familiar faces were clearly distinct from unfamiliar faces. Critically, significant familiarity effects for famous faces were observed in three consecutive experiments, which in turn suggests no special status for personal familiarity. Our conclusion of quantitative differences between different types of familiarity is further supported by relative effect sizes, which, across experiments, gradually decreased from familiarity effects for own (N250 $d_{unb.}$ = 1.04; SFE $d_{unb.}$ = 1.59), personally familiar (N250 $d_{unb.}$ = 0.55; SFE $d_{unb.}$ = 0.92[4]), favourite celebrity (N250 $d_{unb.}$ = 0.40; SFE $d_{unb.}$ = 0.66), disliked celebrity (N250 $d_{unb.}$ = 0.15; SFE $d_{unb.}$ = 0.30) and other celebrity faces (N250 $d_{unb.}$ = 0.27; SFE $d_{unb.}$ = 0.33). While this graded effect is most clearly evident in the SFE, a similar pattern also emerged for N250. However, some of the relevant comparisons in the N250 were not statistically significant and we therefore refrain from making strong claims about this time range.

An interesting question for future research concerns whether the present findings reflect a face-specific mechanism or a more general property of familiarity, in which latter case similar results would be obtained if participants were tested with personally familiar versus famous objects or places. The present results do not speak to the issue of face specificity, and the experiments reported here were not designed to answer such questions. We hope here to contribute to the understanding of how face and person recognition works,

---

[4] Please note that effect sizes for personally familiar and favourite celebrity faces are averaged across experiments (Experiments 1, 2, and 4 for personally familiar faces, Experiments 1, 2, and 3 for favourite celebrities).

independent of whether other stimulus categories are processed in a similar way or not. Future studies may test whether an SFE can be obtained for non-face objects.

Understanding face familiarity is of theoretical as well as potential practical importance (Bauer, 1984; Burton et al., 1999; Ramon & Gobbini, 2018). There are a range of circumstances in which an individual might be motivated to try to conceal their acquaintance with someone; for example in criminal or terrorist investigations. The ERP indices investigated here offer promise in that they are indirect measures (based on an irrelevant butterfly detection task) that do not require explicit recognition and yet show distinct patterns of responses to familiar faces across most individual participants (Wiese, Ingram, et al., 2019; Wiese, Tuttenberg, et al., 2019). The present experiments confirm our previous findings of large effect sizes for personally familiar faces in conventional analyses, and additional bootstrapping analyses show reliable effects for several (but not all) individual participants, especially in the SFE. At the same time, false positive results seem to be highly unlikely in this paradigm (see Experiment 2 in Wiese, Tuttenberg, et al., 2019). Together, these results suggest high sensitivity to detect true familiarity with a facial identity in the absence of an explicit recognition judgment.

In sum, the present results indicate that the familiarity of faces known personally and via the media is not represented in qualitatively different ways. Instead, representations of all types of familiar faces become gradually more robust with increasing familiarity. This principle seems to apply to both visual representations (as reflected in the N250 familiarity effect) and to the integration of person-related identity-specific semantic/episodic information (as reflected in the SFE). We conclude that face representations differ with respect to their degree and not type of familiarity, but that there nonetheless remains a clear and pronounced difference between familiar and unfamiliar faces. Paraphrasing Gertrude Stein, we can conclude that "familiarity is familiarity is familiarity".

References

Ambrus, G. G., Kaiser, D., Cichy, R. M., & Kovacs, G. (2019). The Neural Dynamics of Familiar Face Recognition. *Cereb Cortex, 29*(11), 4775-4784. doi:10.1093/cercor/bhz010

Andrews, S., Burton, A. M., Schweinberger, S. R., & Wiese, H. (2017). Event-related potentials reveal the development of stable face representations from natural variability. *Quarterly Journal of Experimental Psychology, 70*(8), 1620-1632. doi:10.1080/17470218.2016.1195851

Bauer, R. M. (1984). Autonomic recognition of names and faces in prosopagnosia: a neuropsychological application of the Guilty Knowledge Test. *Neuropsychologia, 22*(4), 457-469. doi:10.1016/0028-3932(84)90040-x

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological Studies of Face Perception in Humans. *J Cogn Neurosci, 8*(6), 551-565. doi:10.1162/jocn.1996.8.6.551

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the Self-Assessment Manikin and the Semantic Differential. *J Behav Ther Exp Psychiatry, 25*(1), 49-59. doi:10.1016/0005-7916(94)90063-9

Bredart, S., & Young, A. W. (2004). Self-recognition in everyday life. *Cogn Neuropsychiatry, 9*(3), 183-197. doi:10.1080/13546800344000075

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J. B., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology-Applied, 5*(4), 339-360. doi:10.1037/1076-898x.5.4.339

Bruce, V., & Young, A. (1986). Understanding face recognition. *Br J Psychol, 77 (Pt 3)*, 305-327. doi:10.1111/j.2044-8295.1986.tb02199.x

Burton, A. M. (2013). Why has research in face recognition progressed so slowly? The importance of variability. *Quarterly Journal of Experimental Psychology, 66*(8), 1467-1485. doi:10.1080/17470218.2013.800125

Burton, A. M., Kramer, R. S., Ritchie, K. L., & Jenkins, R. (2016). Identity From Variation: Representations of Faces Derived From Multiple Instances. *Cogn Sci, 40*(1), 202-223. doi:10.1111/cogs.12231

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science, 10*(3), 243-248. doi:10.1111/1467-9280.00144

Butler, D. L., Mattingley, J. B., Cunnington, R., & Suddendorf, T. (2013). Different neural processes accompany self-recognition in photographs across the lifespan: an ERP study using dizygotic twins. *PLoS One, 8*(9), e72586. doi:10.1371/journal.pone.0072586

Caharel, S., Courtay, N., Bernard, C., Lalonde, R., & Rebai, M. (2005). Familiarity and emotional expression influence an early stage of face processing: an electrophysiological study. *Brain Cogn, 59*(1), 96-100. doi:10.1016/j.bandc.2005.05.005

Caharel, S., Jacques, C., d'Arripe, O., Ramon, M., & Rossion, B. (2011). Early electrophysiological correlates of adaptation to personally familiar and unfamiliar faces across viewpoint changes. *Brain Res, 1387*, 85-98. doi:10.1016/j.brainres.2011.02.070

Campbell, A., Louw, R., Michniak, E., & Tanaka, J. W. (2020). Identity-specific neural responses to three categories of face familiarity (own, friend, stranger) using fast periodic visual stimulation. *Neuropsychologia, 141*, 107415. doi:10.1016/j.neuropsychologia.2020.107415

Carbon, C. C. (2008). Famous faces as icons. The illusion of being an expert in the recognition of famous faces. *Perception, 37*(5), 801-806. doi:10.1068/p5789

Clutterbuck, R., & Johnston, R. A. (2002). Exploring levels of face familiarity by using an indirect face-matching measure. *Perception, 31*(8), 985-994. doi:10.1068/p3335

Cumming, G. (2012). *Understanding the New Statistics*. New York: Routledge.

Cumming, G., & Calin-Jageman, R. (2017). *Introduction to the New Statistics: Estimation, Open Science, & Beyond*. New York: Routledge.

Devue, C., & Bredart, S. (2011). The neural correlates of visual self-recognition. *Conscious Cogn, 20*(1), 40-51. doi:10.1016/j.concog.2010.09.007

Di Nocera, F., & Ferlazzo, F. (2000). Resampling approach to statistical inference: bootstrapping from event-related potentials data. *Behav Res Methods Instrum Comput, 32*(1), 111-119. doi:10.3758/bf03200793

Eimer, M. (2011). The Face-Sensitive N170 Component of the Event-Related Brain Potential. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *The Oxford Handbook of Face Perception*. Oxford: Oxford University Press.

Ellis, H. D., & Lewis, M. B. (2001). Capgras delusion: a window on face recognition. *Trends Cogn Sci, 5*(4), 149-156. doi:10.1016/s1364-6613(00)01620-x

Ellis, H. D., Young, A. W., Quayle, A. H., & De Pauw, K. W. (1997). Reduced autonomic responses to faces in Capgras delusion. *Proc Biol Sci, 264*(1384), 1085-1092. doi:10.1098/rspb.1997.0150

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods, 39*(2), 175-191. doi:10.3758/bf03193146

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol Gen, 141*(1), 2-18. doi:10.1037/a0024338

Gallup, G. G. (1970). CHIMPANZEES . SELF-RECOGNITION. *Science, 167*(3914), 86-&. doi:10.1126/science.167.3914.86

Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia, 45*(1), 32-41. doi:10.1016/j.neuropsychologia.2006.04.015

Gobbini, M. I., Leibenluft, E., Santiago, N., & Haxby, J. V. (2004). Social and emotional attachment in the neural representation of faces. *Neuroimage, 22*(4), 1628-1635. doi:10.1016/j.neuroimage.2004.03.049

Gosling, A., & Eimer, M. (2011). An event-related brain potential study of explicit face recognition. *Neuropsychologia, 49*(9), 2736-2745. doi:10.1016/j.neuropsychologia.2011.05.025

Herzmann, G., Schweinberger, S. R., Sommer, W., & Jentzsch, I. (2004). What's special about personally familiar faces? A multimodal approach. *Psychophysiology, 41*(5), 688-701. doi:10.1111/j.1469-8986.2004.00196.x

Hole, G. J., George, P. A., Eaves, K., & Rasek, A. (2002). Effects of geometric distortions on face-recognition performance. *Perception, 31*(10), 1221-1240. doi:10.1068/p3252

Itier, R. J., & Taylor, M. J. (2004). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cereb Cortex, 14*(2), 132-142. doi:10.1093/cercor/bhg111

Jenkins, R., & Burton, A. M. (2011). Stable face representations. *Philos Trans R Soc Lond B Biol Sci, 366*(1571), 1671-1683. doi:10.1098/rstb.2010.0379

Johnston, P., Overell, A., Kaufman, J., Robinson, J., & Young, A. W. (2016). Expectations about person identity modulate the face-sensitive N170. *Cortex, 85*, 54-64. doi:10.1016/j.cortex.2016.10.002

Kaufmann, J. M., Schweinberger, S. R., & Burton, A. M. (2009). N250 ERP correlates of the acquisition of face representations across different images. *J Cogn Neurosci, 21*(4), 625-641. doi:10.1162/jocn.2009.21080

Keenan, J. P., Wheeler, M. A., Gallup, G. G., Jr., & Pascual-Leone, A. (2000). Self-recognition and the right prefrontal cortex. *Trends Cogn Sci, 4*(9), 338-344. doi:10.1016/s1364-6613(00)01521-7

Keyes, H., Brady, N., Reilly, R. B., & Foxe, J. J. (2010). My face or yours? Event-related potential correlates of self-face processing. *Brain Cogn, 72*(2), 244-254. doi:10.1016/j.bandc.2009.09.006

Kramer, R. S. S., Young, A. W., & Burton, A. M. (2018). Understanding face familiarity. *Cognition, 172*, 46-58. doi:10.1016/j.cognition.2017.12.005

Luck, S. J. (2014). *An Introdcution to the Event-Related Potential Technique*. Cambridge: MIT Press.

McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. *Electroencephalogr Clin Neurophysiol, 62*(3), 203-208. doi:10.1016/0168-5597(85)90015-2

Natu, V., & O'Toole, A. J. (2011). The neural processing of familiar and unfamiliar faces: a review and synopsis. *Br J Psychol, 102*(4), 726-747. doi:10.1111/j.2044-8295.2011.02053.x

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113. doi:10.1016/0028-3932(71)90067-4

Ramon, M., & Gobbini, M. I. (2018). Familiarity matters: A review on prioritized processing of personally familiar faces. *Visual Cognition, 26*(3), 179-195. doi:10.1080/13506285.2017.1405134

Rossion, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage, 39*(4), 1959-1979. doi:10.1016/j.neuroimage.2007.10.011

Saavedra, C., Iglesias, J., & Olivares, E. I. (2010). Event-Related Potentials Elicited by the Explicit and Implicit Processing of Familiarity in Faces. *Clinical Eeg and Neuroscience, 41*(1), 24-31. doi:10.1177/155005941004100107

Schweinberger, S. R., & Neumann, M. F. (2016). Repetition effects in human ERPs to faces. *Cortex, 80*, 141-153. doi:10.1016/j.cortex.2015.11.001

Schweinberger, S. R., Pickering, E. C., Jentzsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Brain Res Cogn Brain Res, 14*(3), 398-409. doi:10.1016/s0926-6410(02)00142-8

Sugiura, M., Sassa, Y., Watanabe, J., Akitsuki, Y., Maeda, Y., Matsue, Y., . . . Kawashima, R. (2006). Cortical mechanisms of person representation: recognition of famous and personally familiar names. *Neuroimage, 31*(2), 853-860. doi:10.1016/j.neuroimage.2006.01.002

Tanaka, J. W., Curran, T., Porterfield, A. L., & Collins, D. (2006). Activation of preexisting and acquired face representations: the N250 event-related potential as an index of face familiarity. *J Cogn Neurosci, 18*(9), 1488-1497. doi:10.1162/jocn.2006.18.9.1488

Tranel, D., Damasio, H., & Damasio, A. R. (1995). Double Dissociation between Overt and Covert Face Recognition. *J Cogn Neurosci, 7*(4), 425-432. doi:10.1162/jocn.1995.7.4.425

Urbach, T. P., & Kutas, M. (2002). The intractability of scaling scalp distributions to infer neuroelectric sources. *Psychophysiology, 39*(6), 791-808. doi:10.1111/1469-8986.3960791

Wiese, H., Ingram, B. T., Elley, M. L., Tuttenberg, S. C., Burton, A. M., & Young, A. W. (2019). Later but not early stages of familiar face recognition depend strongly on attentional resources: Evidence from event-related brain potentials. *Cortex, 120*, 147-158. doi:10.1016/j.cortex.2019.06.004

Wiese, H., Tuttenberg, S. C., Ingram, B. T., Chan, C. Y. X., Gurbuz, Z., Burton, A. M., & Young, A. W. (2019). A Robust Neural Index of High Face Familiarity. *Psychological Science, 30*(2), 261-272. doi:10.1177/0956797618813572

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: the SHINE toolbox. *Behav Res Methods, 42*(3), 671-684. doi:10.3758/BRM.42.3.671

Young, A. W., & Burton, A. M. (2017). Recognizing Faces. *Current Directions in Psychological Science, 26*(3), 212-217. doi:10.1177/0963721416688114

Young, A. W., & Burton, A. M. (2018). Are We Face Experts? *Trends Cogn Sci, 22*(2), 100-110. doi:10.1016/j.tics.2017.11.007

Zimmermann, F. G. S., Yan, X., & Rossion, B. (2019). An objective, sensitive and ecologically valid neural measure of rapid human individual face recognition. *R Soc Open Sci, 6*(6), 181904. doi:10.1098/rsos.181904

Appendix 1, celebrity images in Figure 1

Sandra Bullock 01
Source: https://www.flickr.com/photos/evarinaldiphotography/9192365016/
Attribution: Eva Rinaldi
Licensed for reuse under the Creative Commons Attribution-Share Alike 2.0 Generic license
https://creativecommons.org/licenses/by-sa/2.0/deed.en

Sandra Bullock 02
Source: https://www.flickr.com/photos/evarinaldiphotography/9189702847/
Attribution: Eva Rinaldi
Licensed for reuse under the Creative Commons Attribution-Share Alike 2.0 Generic license
https://creativecommons.org/licenses/by-sa/2.0/deed.en

Sandra Bullock 03
Source: https://www.youtube.com/watch?v=nkSu6RCaF7g
Attribution: MTV International https://www.youtube.com/user/MTVUKofficial
Licensed for reuse under the Creative Commons Attribution 3.0 Unported license
https://creativecommons.org/licenses/by/3.0/deed.en

Sandra Bullock 04
Source:
https://commons.wikimedia.org/wiki/File:Sandra_Bullock,_The_Heat,_London,_2013.jpg
Attribution: Richard Goldshmidt
Licensed for reuse under the Creative Commons Attribution 3.0 Unported license
https://creativecommons.org/licenses/by/3.0/deed.en

Sandra Bullock 05
Source: https://www.flickr.com/photos/22007612@N05/9354456681
Attribution: Gage Skidmore
Licensed for reuse under the Creative Commons Attribution-Share Alike 2.0 Generic license
https://creativecommons.org/licenses/by-sa/2.0/deed.en

Sandra Bullock 06
Source: https://www.flickr.com/photos/viatorci/5492390318/
Attribution: David Torcivia
Licensed for reuse under the Creative Commons Attribution-Share Alike 2.0 Generic license
https://creativecommons.org/licenses/by-sa/2.0/deed.en

Sandra Bullock 07
Source: https://www.flickr.com/photos/tomdog/4299833829/
Attribution: djtomdog https://www.flickr.com/people/51761894@N00
Licensed for reuse under the Creative Commons Attribution-Share Alike 2.0 Generic license
https://creativecommons.org/licenses/by-sa/2.0/deed.en

Sandra Bullock 08
Source: Own work
Attribution: Georges Biard

Sandra Bullock 09
Source: https://www.flickr.com/photos/evarinaldiphotography/9189571805/
Attribution: Eva Rinaldi

Sandra Bullock 10
Source: https://vimeo.com/275657737
Attribution: Eva Rinaldi

Kate Ritchie
Source: https://www.flickr.com/photos/evarinaldiphotography/8784938666/
Attribution: Eva Rinaldi